RJ10513 (ALM1309-013) September 23, 2013
Computer Science

IBM Research Report

Identifying User Needs from Social Media

Huahai Yang, Yunyao Li
IBM Research Division
Almaden Research Center
650 Harry Road
San Jose, CA 95120-6099
USA

Identifying User Needs from Social Media

Huahai Yang
IBM Research - Almaden
650 Harry Rd. San Jose, CA 95120
hyang@us.ibm.com

Yunyao Li
IBM Research - Almaden
650 Harry Rd. San Jose, CA 95120
yunyaoli@us.ibm.com

ABSTRACT

With the rise of social media, writings by ordinary people are becoming increasingly available for linguistic analysis. Such analyses offer great opportunities to identify individual users' needs from user-generated content, so that better tailored products or services can be recommended. Literature suggests that several types of human needs are universal and directly influence consumer purchase behavior. In this paper, we investigate the use of social media to identify such fundamental needs for individuals. We developed psychometric measures of universal needs through a crowd-sourced study. We also built several models to predict people's needs based on their writings. We conducted a detailed analysis of the models and showed that our models can effectively identify users' needs based on their social media data. Our results also confirm that some inferred needs correlate well with the actual product purchases and suggest a great potential for our models to significantly increase effectiveness of product recommendations.

Author Keywords

Needs; social media; psychometrics; natural language processing.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

The rise of social media, as exemplified by Facebook and Twitter, has led to an abundance of user-generated content on the Web. The permeation of social media technology into people's daily lives has created unprecedented opportunities for researchers to link linguistic data with all aspects of human behaviors in order to better understand people. Many aspects of human behaviors have been explored using social media data, such as detecting and monitoring disaster responses [37], forecasting movie success [4] and characterizing debate performance [14], among others.

One important aspect of human behavior yet to be investigated in the context of social media is people's needs[15, 23]. One great potential for research on needs is to help product/service providers, such as businesses, government, or non-profit institutions, to better understand their customers and thus improve their offerings to the general public. For example, such research may enhance the quality of direct marketing, which targets selected population with tailored messages [6]. Instead of marketing indiscriminately to the masses, personalized information about selected products or services could be promoted to people in need of them. Such targeted marketing is not only likely to increase the success rate of marketing campaigns but also reduce annoyance and discomfort on the part of the consumers. In fact, since needs fulfilment is known to be associated with enhanced well-being [42], the research effort of inferring needs from social media could positively contribute to improving people's quality of life.

The goal of this work is to build predictive models that can automatically infer people's fundamental needs from user-generated content. Our contributions towards this goal are summarized as follows: (1) We developed psychometric measures of universal needs via crowd-sourcing; (2) We collected people's writings about their needs from the study and used the data to train several models to identify people's needs scores from their writings; (3) We conducted a detailed analysis of the models in simulated direct marking scenarios. The results show that our models can effectively identify some fundamental needs from social media. The results also confirm that some of the inferred needs correlate very well with actual product purchases. Implications and future work are discussed at the end of the paper.

RELATED WORK

This work draws from the research areas of psychology, psychometrics, marketing, and natural language processing.

Universal Needs and Consumer Behavior

First, we clarify that the kind of needs we study is not merely of the "I want an iPad" type. Such specific desires for concrete things are contextual manifestation of deeper needs, which Maslow postulates as a hierarchy of fundamental and universal needs [25, 26],[1]. He suggests that human behaviors are motivated by expression and fulfilment of these deeply held needs, which people may not be entirely consciously aware of. The hierarchical and ordering properties of these needs have been found to be weak [45]. However, the existence of universal needs has been recently reaffirmed and is found to be connected with people's subjective well-being [35, 36, 42]. We also observe that the current explosion

[1]Namely, *physiological*, *safety*, *love/belonging*, *esteem*, *cognition*, *aesthetics*, and *self-actualisation*

of social media data may help solve what Maslow lamented as a "very serious lack of sound data" [25] problem in the research on fundamental needs. Our work could be considered as such an effort.

To help establish external validity of our research on human needs, we need to pick an application area that fundamental needs have obvious and strong influence. In this paper, we chose to work with the area of marketing research. One reason is that universal needs are already explicitly invoked to explain consumer behaviors such as brand association and customer satisfaction [15, 23]. Though similar to Maslow's theory, the classification of needs in the marketing context may use different labels and have different number of categories. These categories often originate from consumer surveys, where consumers rate or compare products on various attributes. Dimension-reduction statistical methods, such as multidimensional scaling, multiple discriminant analysis, or principle component analysis, are then used to produce diagrammatic representation of consumer perceptions— so called *perceptual maps* [20, 22]. Ford [15] observes that these perceptual maps often align well with systems of universal needs. However, perhaps due to Maslow theory's being out-of-fashion for many years [42], unlike other personal traits such as personality [19] or values [39], standard psychometric scales have not been developed to directly measure people's fundamental needs. Our work fills the void.

Building Predictive Models from User-Generated Content

The prevalence of social media content leads to increasing interests in building predictive models for a wide range of topics through linguistic analysis over social media. One important direction is personal profile prediction based on user-generated content. Much work in this direction build on top of the existing research on sociolinguistics and focus on automatically identifying latent demographic features of online users, including gender, age, education, religions, and political orientation, among others [2, 7, 28, 34, 38]. Another line of related work in this direction seeks to reveal the personality traits of online users through their use of social media [17, 18, 33]. It has been shown that personal traits can be predicted with relatively high accuracy by analyzing public data that people share online.

As mentioned earlier in the introduction, another widely explored research direction is building predictive models for human behaviors. Work in this direction include building models to detect and monitor disaster responses [37], forecast movie success [4] and election results [40, 43], characterize debate performance [14], and predict depression [11], among others. One human behavior of particular interest to our work is purchase behavior. Zhang et al. [47] built predictive models on purchase behavior using users' Facebook profile information. Site visitors' activities on the e-commerce site is used to predict buying propensity [41], so is the usability of the site itself [44].

While related, our work fundamentally differs from previous approaches: (1) we built our predictive models for user needs (which in turn map to certain purchase behavior) solely based on user-generated content; (2) the purchase behavior under consideration is not limited to one specific e-commerce site, but all of the user's consuming venues, in so far as these purchases are revealed on social media; and (3) the predicted purchase is not limited to the product categories that the model has been trained on (e.g. either cars, electronics, or fashion), but runs cross many product categories relevant to one need (i.e. someone may buy luxury goods from a diverse set of product categories to fulfil the same need for **ideal**).

TRAINING DATA COLLECTION

To achieve the goal of predicting people's fundamental needs from social media, a significant amount of training data with ground truth is necessary. The ideal method of data collection would be to ask a large number of social media users to take a standard psychometric test of fundamental needs. Unfortunately, this method is not feasible in our case. The first obstacle is the aforementioned lack of standard psychometric measure for fundamental needs. We attack this problem heads-on by developing our own psychometric measures based on existing literature. Secondly, making unsolicited requests to strangers is against the terms of services of most of the social media sites. Finally, people are increasingly concerned about privacy issues on the Internet. Therefore, we have made respecting users' privacy the highest priority. Personal identifiable information are not collected anywhere in our research process.

Develop Psychometric Measure of Needs

In Maslow tradition, basic needs are considered innate. However, the expression and fulfilment of them are not constant in one's life. The hierarchy implies a priority among the needs. At a given point in time, some of a person's fundamental needs are more prominently featured, while other needs are "minimized" [25]. It therefore makes sense to have a notion of a person's *current* state of fundamental needs. The goal of psychometric measure of needs is to take a snapshot of the current state.

In order to develop such a psychometric scale of needs, we first need to decide among the existing systems of needs classification[13]. Fortunately, newer classifications are not fundamentally different from Maslow's original proposal. For our purpose, we chose to use a marketing practice oriented classification [15]. This system is more finely delineated than other systems, containing the following 12 needs: **structure**, **practicality**, **challenge**, **self-expression**, **excitement**, **curiosity**, **liberty**, **ideal**, **harmony**, **love**, **closeness** and **stability**. More importantly, each need in this system is clearly associated with matching properties of the desirable products and services. Table 3 shows some example associations spelled out by the book. As described later in this paper, these clear definitions help us to design experiments to examine the validity of the needs measures.

We followed the standard practice for psychometric scale development [3, 8]. For each needs type, we initially developed 10 question items based on the description. An example questions item for the *ideal* is "*I crave refinement in life.*" (see Appendix B for current question items for *ideal* and *closeness*). Survey respondents rated their agreement with the statement

Example 1: `1. I need to pay off all my credit card debt by next year. I'm cutting cost wherever I can to help pay off these debts. This will help get me more financially stable. 2. I want to take a vacation to Florida. This also requires saving and cutting costs. I haven't had a vacation in a long time. 3. I want to lose weight and get healthy. There are sports that I would like to get back into and can't with my current weight.`

Example 2: `1. I want to go running again. I went through a lot of trauma a year ago, and it was reinforced by a couple of troubling incidents that have happened more recently. I find that I can't go out without being paranoid about who's around and anticipating that I'll get attacked and humiliated again. But I so miss the freedom of running. I miss moving my body, my limbs, and my quick responses to challenges that pop into my head when I'm running. I want to get back into it. 2. I need to learn to trust people. I think that everyone has an ulterior motive or is just using me for favors I'll do for them. These thoughts pop into my head when people are making sincere confessions of emotion, and I still don't trust them. I would have more friends if I trusted people more. It's good to be cautious and skeptical, but skepticism needs to be rational. 3. I want to start some kind of company that serves my community. Hair, clothing, whatever. I think it would be fun and would be an outlet for me to 1) meet people and 2) express my creativity and love for fashion.`

Figure 1: Sample written responses

using a five-point Likert scale, with 0 being "*not at all*" and 4 being "*very much*". Respondents were also asked to think of things they "*recently sought after or are considering to get*" and describe their current needs, since needs may change as the person's circumstances change.

We used Mechanical Turk to expedite the scale development. Care was taken to ensure the quality of the responses: only US workers were allowed to participate in order to reduce the noise of cultural factors[2]; participants were required to have more than 99% work acceptance rate; one person could respond only once; only completed surveys were accepted; each response was individually screened before approval; substandard survey responses were removed[3]; and the respondents were paid substantially higher wages than the Mechanical Turk average.

The initial item pool was tested by 360 respondents. We plotted and examined the response distribution of each need and eliminated question items that produce highly skewed distribution exhibiting ceiling or flooring effect. In the end, only those question items producing a broad distribution were retained. The second round of tests used another 360 respondents. The responses were then analysed for structural consistency. Items were screened to ensure the inter-item correlation for a needs category falls in the range of [0.15, 0.5]: too high a correlation indicates redundant questions, while too low a correlation suggests inconsistency. For needs categories that already had less than four questions, wordings of the question statement were adjusted to move the distribution of the responses to the desired direction. After the third round of tests, we obtained approximately normal distributions of the averaged scores for most of the needs types (except for *curiosity*, which was still strongly skewed to the positive side). The subsequent large scale data collection used this version of the survey, which consisted of 55 questions in total, with 4 or 5 question per need type. As a method to catch inattentive respondents, one question was repeated once in a distance location of the survey. A total of 2587 approved survey responses to this version of the survey were received in the end.

Collect User-Generated Content
Since we decided not to collect personally identifiable information, we did not crawl respondents' social media writings directly. Instead, we asked our anonymous survey respondents to write about their needs. The instruction reads: "*Please describe three things that you want to get or need to do the most, and explain why you want or need them. Please be as honest as possible*". Minimal 60 words were required for the answer. On average, each respondent wrote 103 words (see Figure 1 for sample responses). The responses were collected as the training data for building our needs models.

NEEDS MODELS
After collecting a large number of survey responses, we proceed to train machine learning models to predict an author's needs score given his writings.

Features
While metadata (e.g. timestamps, geo-locations and user profiles) associated with social media can be useful, to build discriminative needs models that work for any user-generated text content, we only consider features that occur within the main body of user-generated textual content (e.g. Tweets). Moreover, since our models are trained on survey responses that are very different from social media text, we seek to build models that are highly context-independent and robust across multiple content domains. Simple word lists such as LIWC [31] have shown to be revealing and versatile features for building models that associate the way people write with their deep personal traits [30, 46]. While the word lists in LIWC were produced through brainstorming by experts, our goal was to automatically generate lists of words that are indicative of the presence or absence of fundamental needs. The following set of linguistic features were considered:

[2] We plan to develop versions of the survey in other countries.

[3] Such as giving different answers to a pair of identical question appearing in different parts of the survey; responses taking too little time to complete; and responses with the same choices consecutively for more than 10 questions

[4] "***" indicates that the correlation is statistically significant, $p < .001$; "N/A" indicates that the algorithm fails to obtain any non-zero coefficient for any of the features.

Need	Correlation		Mean Absolute Error	
	TEXTBASIC / TEXTEXPANDED	TEXTFILTERED / TEXTEXPANDEDFILTERED	TEXTBASIC / TEXTEXPANDED	TEXTFILTERED / TEXTEXPANDEDFILTERED
structure	.09***	N/A	.53	.53
practicality	.24***	N/A	.49	.50
challenge	.39***	.18***	.63	.65
self-expression	N/A	.15***	.61	.61
excitement	.48***	.44***	.71	.73
curiosity	N/A	N/A	.57	.57
liberty	.11***	.24***	.55	.54
ideal	.24***	.27***	.68	.68
harmony	.20***	N/A	.51	.51
love	.25***	.20***	.68	.69
closeness	.32***	.17***	.61	.63
stability	N/A	N/A	.53	.53

Table 1: Results of 10-fold cross validation [4]

closeness		**challenge**		**excitement**		**love**		**ideal**	
family	1.196	final	0.353	college	0.859	love	0.560	love	0.349
love	0.557	career	0.199	world	0.599	friend	0.680	precious	0.087
your	-0.111	marathon	0.176	gym	0.195	breakdown	-0.379	pick	-0.014
critical	-0.928	design	-0.976	litterbox	-0.001	item	-0.408	bottle	-0.163
handmade	-1.577	attack	-0.218	doctor	-0.018	state	-0.586	attack	-0.676

Table 2: A sample set of features from TEXTBASIC for different needs.

I. TEXTBASIC: Unique unigrams from the training data. We used normalized term frequency for unique unigrams [5]. We removed all unigrams that contain non-word characters or digits. Stopwords were retained as they can be informative (e.g. *we* correlates with **closeness**). No stemming was performed as stemming can be lossy (e.g. *family* correlates with **closeness** while *families* does not.)

II. TEXTEXPANDED: Hoping to increase the coverage of the models, we expanded features in TEXTBASIC with the synonyms from the unique grams based on WordNet [27]. When multiple synsets existed for the same unigram, we chose all synonyms from the top ranked synset. We did not add synonyms that already present in TEXTBASIC.

III. TEXTFILTERED: As a potential way to reduce noise, we filtered features in TEXTBASIC by keeping only those that occurred in sentences containing a short list of verbs indicating needs (i.e. *need*, *want*, and *wish*). We used Stanford Dependency Parser [12] to identify such verbs.

IV. TEXTEXPANDEDFILTERED: Similar to TEXTFILTERED, we filtered features in TEXTEXPANDED by keeping only those that occurred in sentences containing the verbs indicating needs.

To predict measured needs scores using the above features, we chose the generalized linear regression model with L_1 regularization, which does both coefficients shrinkage and selection [16]. We had also tried Support Vector Machine(SVM) [9] for model training. However, SVM models produced worse R-squared measures than lasso models and thus are omitted in the rest of paper. Each of the feature sets above generates one model. For simplicity, we refer to each model by the name of its corresponding feature, unless noted otherwise. The generated models all contain a small number of words with non-zero coefficients.

Model Fit

We performed 10-fold cross validation during model training. Table 1 reports the correlations between the predicted and measured needs scores and Mean Absolute Error (MAE) for all four models. Note that the statistics for TEXTEXPANDED and TEXTEXPANDEDFILTERED are the same as those for their corresponding non-expanded models, TEXTBASIC and TEXTFILTERED, respectively. The reason is that expanded words in these two models do not appear in the original training data and thus have no impact on MAEs. As can be seen, we have successfully built language feature based models that significantly correlate with 10 out of the 12 measured needs. Our best performing models are for **excitement**. On the other end, two measured needs completely escaped our effort to capture them with models. For **curiosity**, it might due to our failure to obtain a properly shaped item response distribution. Better feature might be necessary for **stability**. All of the MAEs are not very small, indicating that there is still room for improvement for future needs modelling effort.

Table 2 lists a selected sample of words learned and their weights for TEXTBASIC. As can be seen, most of the features do make intuitive sense, while a few of them are counter-intuitive. For instance, it is easy for one to understand why "*family*" and "*love*" are positively associated **closeness**, while "*your*" and "*critical*" carry negative weights. However, it is harder to explain why "*handmade*" also carries a negative weight. Similarly, it is natural for one to associate '*college*", "*world*", and "*gym*" positively with **excitement**, and "*litterbox*" and "*doctor*" the opposite way. In addition, the same feature may occur in more than one need. For instance, "*love*" carries positive weights for **closeness**, **love**, and **ideal**, while "*attack*" carries negative weights for both **challenge** and **ideal**. This observation indicates that the needs are not completely independent, consistent with what has been suggested by the literature [26].

EXPERIMENTS

To evaluate the validity of our needs psychometric scales and the effectiveness of our needs models, we designed a set of experiments to simulate direct marketing scenarios. The goal is twofold: (1) to verify whether the purchase behavior of Twitter users are indeed consistent with what universal needs would prescribe and what our needs models would predict; (2) to evaluate the usefulness of our approach of building needs models via survey to predict purchase behavior by comparing with a prediction model directly trained with social media data.

Dataset

In order to validate whether our model can indeed predict users' needs based on their social media content, we first need to collect the ground truth. In another words, for each twitter user whose tweets we are analyzing, we need to know his/her actual needs. The ideal approach is to contact Twitter users and ask them to fill out our survey, from which we can establish the ground truth about their needs. Unfortunately, as discussed earlier this approach is not feasible in our case.

Instead, we took the alternative approach to establish user needs as ground truth, by searching certain information explicitly revealed by their tweets. Specifically, we searched and identified those tweets containing mentions of specific product purchases that match certain needs as prescribed by the marketing literature [15]. To do so, for each need of interest, we first collected top sellers from the appropriate product categories on Amazon.com[6]. We then used carefully handcrafted patterns (e.g. "`I got` $\langle ProductName \rangle$") to identify purchases of these products. Since multiple product categories are mentioned for each of the 12 needs types, with each product category associated with many products, it is impractical to test of all the needs in one experiment.

For the purpose of this paper, we chose to test only two of the needs, **closeness** and **ideal**. These two needs were chosen for their importance to the most common sales strategies: the former need for cross selling[7] and the latter for upselling[8] [15]. We believe that our choice best serves the main goal of the study — to experimentally validate the value of needs prediction in a practical setting, even though our models for these two needs do not perform the best. Conducting similar experiments for the remaining needs is currently undergoing but is outside of the scope of this paper.

The types of product purchases identified from tweets are listed in Table 3. This process was done using information extraction with highly precise textual patterns (e.g. "`I just bought` $\langle ProductName \rangle$") over the tweets (excluding re-tweets) generated in December 2012. Based on the Twitter ids associated with the extracted tweets, we identified two group of Twitter users: one group with **closeness**, the other with **ideal**. We then randomly selected 100 users from each group and collected all the tweets of these users since 6 months before and during December 2012. In addition, we randomly crawled tweets of another 900 Twitter users not identified with either needs based on information extraction. This process resulted in two collections of tweets, each with 1000 unique Twitter users: one for evaluating **closeness**, referred to as $\text{TWEETS}_{closeness}$, and the other for evaluating **ideal**, referred as TWEETS_{ideal}. The average number of tweets collected per user is 2685 for $\text{TWEETS}_{closeness}$ and 2441 for TWEETS_{ideal}. Both datasets were used as our testing data, consisting of over six million tweets in total.[9]

Measures

In predicting purchase behavior, socio-economic variables, demographics information, attitudes, and buying intentions have all been found to be less satisfactory than the measured subjective purchase probability [10]. As discussed before, it is unrealistic to ask a large number of social media users to fill out purchase probability surveys. Therefore, we adopted *Lift Index* [24], a common measure for evaluating predictive models in direction marketing, to assess and compare our models.

To use the *Lift Index*, a predictive model must generate a ranked list for all the test samples according to its prediction. Given such a ranked list of test samples S, the lift index for the model is calculated as

$$Lift = (1 \times S_1 + 0.9 \times S_2 + ... + 0.1 \times S_{10}) / \sum_{10}^{i=1} S_i \quad (1)$$

with S_i denoting the number of actual buyers in the top ith decile of the ranked list. The *Lift Index* measures the power of the model to move high potential buyers to the top ranked positions. An optimal predictive model would produce the lift index value of 100%, while a completely random model would result in a lift index of 50% [24].

[5] Certain content (indicated by "[***]") has been omitted to fit the tweets in the table.

[6] Based on **`http://www.amazon.com/Best-Sellers/zgbs`** in June 2013

[7] Convince a customer to buy additional product that is very different from the existing purchase.

[8] Convince a customer to buy more expensive product.

[9] We plan to release the datasets for research purpose.

Need	Type of Product Purchases	Example Tweets
closeness	Period TV Series	`I'm watching That '70s Show (16 others checked-in) [***]`
	Home Decoration	`I just bought skittle candles`
	Home and Garden Magazine	`I LOVE Atomic Ranch magazine!` **`http://t.co/rP04SQjE`**`!`
ideal	Organic Skin Care Products	`I just bought: 'Burt's Bees Baby Bee Getting Started Kit' [***]`
	Organic Food	`I bought organic milk today. Just thought you would be proud [***]`
	Pet Clothing	`Just bought my dog a coat for her christmas present #socute`

Table 3: Sample types of product purchases and example tweets[5]

	ideal	**closeness**
TEXTBASIC	58.9%	58.9%
TEXTEXPANDED	59.4%	58.9%
TEXTFILTERED	70.5%	**62.5%**
TEXTEXPANDEDFILTERED	**74.5%**	62.2%
TEXTTWITTER	N/A	60.0%

Table 4: Lift index of all models

Specifically, we use our needs models to score social media users according to their likelihood to have a specific need under consideration. Since each need strongly correlates with specific product purchases(e.g. Table 3), we can rank the users according to their need score from most likely to least likely buyers. The lift reflects the redistribution of social media users after the ranking, with superior model showing a high concentration of actual buyers in the upper quantiles of the ranked list. Hence, the lift evaluates a model's capability to identify potential buyers' needs that are strongly associated with certain types of product purchases and measures the improvement over selecting customers for marketing at random.

Note that in our datasets the actual buyers were determined based on information extraction over tweets. While the information extraction algorithm is highly precise, it is possible that some of the actual buyers were mislabeled as non-buyers due to (1) the limitation of our information extraction process; and/or (2) actual product purchases were not mentioned in the tweets. As a result, our models may be wrongly penalized for correctly identifying those mislabeled users as actual buyers. Therefore, it is worth noting that the lift index only reflects the lower bound of our models' capability.

Effectiveness of Needs Models

We first evaluate the effectiveness of our needs models by examining whether the purchase behavior of Twitter users is indeed consistent with what universal needs would prescribe and what our needs models would predict. We applied all four needs models discussed earlier over our testing data TWEETS$_{closeness}$ and TWEETS$_{ideal}$. Table 4 reports the lift index for all models (TEXTTWITTER is described later when we discuss the usefulness of needs modeling). As can be seen, the lift index scores of all four models are higher than 50% for both **ideal** and **closeness**. This result suggests that our needs models can indeed effectively identify universal needs and that such needs are indeed consistent with the purchase behavior of Twitter users.

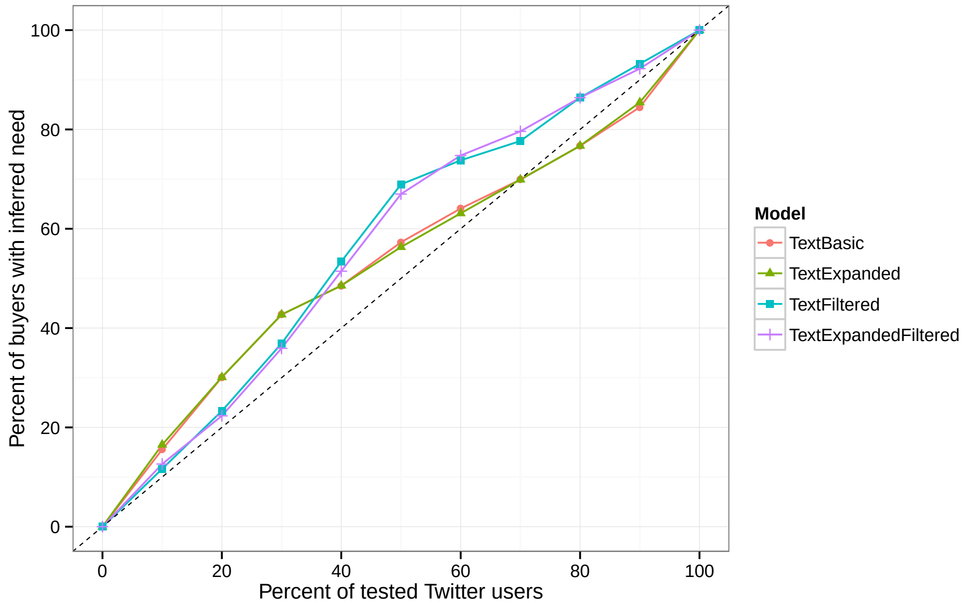
In addition, we can see that models using filtered features, TEXTFILTERED and TEXTEXPANDEDFILTERED, are clear winners — both perform more effectively than the models without filtering. This performance advantages of models with filtered features indicate that the simple verb-based filtering method does have helped removing some of the noisy features. We can also observe that models with features expanded using WordNet expansion generally perform slightly better than the ones without, with the exception of one case: the lift index of TEXTFILTERED for **closeness** is slightly higher than that of TEXTEXPANDEDFILTERED. This minuscule discrepancy suggests that WordNet expansion may have introduced a small amount of noise into the model. Nonetheless the best overall model is arguably TEXTEXPANDEDFILTERED, the model with both feature filtering and expansion.

Besides evaluating the overall effectiveness of the models, we are also interested in comparing the effectiveness of the models in terms of helping "skim the cream" by selecting a relatively small number of the Twitter users for a relatively large portion of the actual buyers. This comparison can be done visually by examining the lift charts. Figures 2a and 2b show the cumulative lift curves for all models for **ideal** and **closeness** respectively. As one can see, all four models almost always consistently outperform the random, with the entire or most of the curves above the diagonal line. While TEXTEXPANDEDFILTERED remains superior than other models for **ideal**, regardless of the cut-off point, it is no longer the clear winner for **closeness**. In fact, for lower cut-off points (e.g. 10%), the baseline models without feature filtering actually provide higher lift. This finding indicates that the simple feature filtering we performed, although is helpful in improving the overall performance of the models, may have removed some useful features for certain needs.

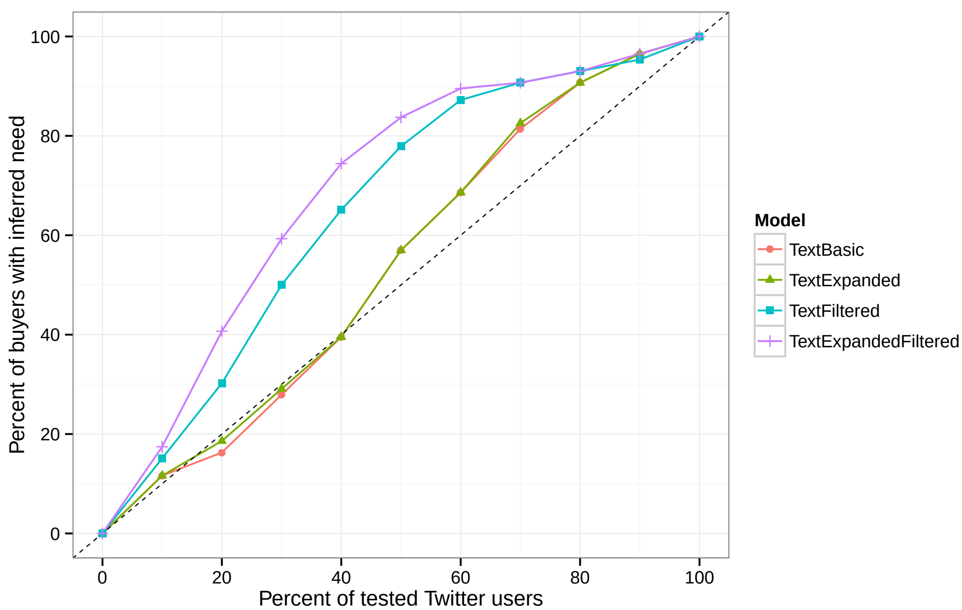
Usefulness of Needs Modeling

Results above suggest that our needs modeling approach can be effective in predicting purchasing behavior. It is of interest to see how it compares with other solutions, especially models trained directly for the task using the dataset of interest.

We crawled the tweets of 1000 random Twitter users as negative samples, and for the positive samples, we collected the tweets of 100 additional Twitter users who purchased products that associate with **closeness** but not in TWEETS$_{closeness}$. We then used this dataset as the training data to build a predictive model using unigram features similar to TEXTBASIC, but based on tweets instead of survey responses. For the ease of comparison, we used the same



(a) For **closeness**



(b) For **ideal**

Figure 2: Cumulative lift curves (Dashed diagonal line indicates a random model)

generalized linear regression model with L_1 regularization described earlier. The resulting model is referred to as TEXTTWITTER.

We applied TEXTTWITTER to TWEETS$_{closeness}$ and obtained a lift index of 60.0%. As can be seen from Table 4, the lift index of TEXTTWITTER, while evidently lower than that of our best models, is slightly higher than that of both TEXTBASIC and TEXTEXPANDED. This result confirms that our needs modeling approach is useful. However, the slight edge of TEXTTWITTER over TEXTBASIC and TEXTEXPANDED suggests that additional features may have been learned from the raw social media data that are not available in the survey data. While enhancing our needs models using features learned directly from social media is a promising direction to explore, it is out of the scope of the current paper.

DISCUSSIONS

This work contributes to HCI by making connections with a number of related fields.

Measuring Needs

As far as we are aware of, we are the first to develop a survey instrument for fundamental needs. As such, the instrument still requires more work to prove its reliability and validity. The experiments reported here certainly demonstrate certain external validity for two of the needs categories. On the other hand, the transient nature of needs makes test-retest reliability less practical to do since we are taking a snapshot of needs. The temporal change of needs would be an interesting research topic. We also see that the items for some needs are better than others. For example, we have a hard time developing good items for the need of **curiosity**. We plan to continue the work to refine the psychometric properties of the needs scale, and make the instrument available for others to test.

As a basic psychometric instrument, we expect the needs scale to be useful in many fields of research and application. For example, based on the strong connection between needs and human emotion [25], the survey could be used as a research tool in affective computing [32]. It would be an attractive proposition to built an intelligent user interface sensitive to people's fundamental needs. On the other hand, we showed that our methodology of crowd-sourced psychometric approach can be very effective. This methodology can be adopted to test other personal traits of interest (e.g. *friendliness*).

Modelling Needs

Our experimental results have confirmed that our needs models can effectively predict purchase behaviors. As discussed earlier, the models we built are simple word lists similar to LIWC. Such models, although simple, are surprisingly powerful. The simplicity of the models not only permits them to be easily applied to different types of user-generated content, but also make them extremely flexible to use — the models can be applied by themselves, or used as part of more sophisticated learning models (e.g. one that takes social media profile and structure into consideration). The latter is an interesting research direction to explore in the future. We discovered a few potential weakness of the current models in feature filtering and expansion. We plan to continue work towards improving our model by exploring ways to enhance the feature filtering and expansion.

From an application point of view, one may argue that instead of inferring universal needs for social media users, it may be more effective to directly infer their intention to buy a product. Indeed, identifying purchase intent from social media is an emerging topic by itself [1, 29, 21]. However, identifying sales lead in the form of "*intention to buy* $\langle ProductName \rangle$" is also widely regarded as a "needles in a haystack" problem [1]. In another word, only a tiny fraction of social media users can be associated with any specific purchase intent. We have experienced this scarcity while collecting the data set for our experiments. In contrast, universal needs are considered as fundamental intrinsic properties associated with each individual. Like most such psychological properties, the distribution among the general population is close to be Gaussian. Using universal needs as the surrogates for purchasing behaviors suffers less from the sparse signal problem, as confirmed by our experimental results.

We also would like to further evaluate the value of the needs model in real direct marketing scenarios. In addition, we are interested in applying the needs models to predict other human behaviors. For example, we may use the models to evaluate employees' well-being for employee retention program, help matching mentors and mentees, and improve online dating services by better matching candidates. Another interesting future direction to pursue is to investigate how people's needs change over time and how such changes correlate with specific life events (e.g. graduation, marriage, divorce and so on).

Limitation and Future Work

In our current experiments, for each need we evaluated our models over multiple products at the same time. Intuitively, some products may be better predicted than others by the same need. In addition, one product may be better predicted by multiple needs. If we want to develop the needs prediction task into a new standard natural language processing task on par with sentiment analysis, we need to define the task more rigorously and to develop standard data sets and corresponding ground truth. In this regard, this task is significantly different from sentiment analysis, in that a person without training in psychology can label sentiment if he/she can read, but the same person may have a hard time labelling needs by reading alone. In this paper we showed one possible approach to gather ground truth through crowd-sourcing, developing other methods would also be welcomed.

We observe that a group of less successful modelled needs are closely related: **stability**, **practicality**and **structure**. Future research should study the reason. Better modeling approach using semantic analysis might help. However, an alternative hypothesis is that this is a culture artefact: people in Western culture may be discouraged from expressing needs placed low in Maslow's hierarchy. Similarly, we found that **curiosity**is very hard to measure using self reporting survey. Is this also a Western culture peculiarity? We doubt it, but measuring and modelling needs in another culture would definitely make a fascinating cross-culture research topic.

CONCLUSION

We have developed psychometric measures of universal needs via a crowd-sourced study. We have also explored the linguistic relationships between people's fundamental needs and their social media data. Through a detailed analysis of the models, we show that the our models, although simple, are able to predict people's fundamental needs that are strongly correlate with actual product purchases. More broadly, our work is an example of using crowd-sourced studies to develop psychometric measures and predictive models for human behaviors.

REFERENCES

1. Alexe, B., Hernandez, M. A., Hildrum, K. W., Krishnamurthy, R., Koutrika, G., Nagarajan, M.,

Roitman, H., Shmueli-Scheuer, M., Stanoi, I. R., Venkatramani, C., et al. Surfacing time-critical insights from social media. In *SIGMOD* (2012), 657–660.

2. Argamon, S., Koppel, M., Pennebaker, J. W., and Schler, J. Mining the blogosphere: Age, gender, and the varieties of self-expression. *First Monday 12*, 9 (2007).
3. Association, A. P. *Standards for Educational and Psychological Testing*. American Educational Research Association, 1999.
4. Asur, S., and Huberman, B. A. Predicting the future with social media. In *Proceedings of Web Intelligence and Intelligent Agent Technology (WI-IAT)* (2010), 492–499.
5. Baeza-Yates, R., Ribeiro-Neto, B., et al. *Modern information retrieval*, vol. 463. ACM press New York, 1999.
6. Bob, S. *Successful direct marketing methods*. NTC Business Books, 1989.
7. Burger, J. D., Henderson, J., Kim, G., and Zarrella, G. Discriminating gender on Twitter. In *EMNLP* (2011), 1301–1309.
8. Clark, L. A., and Watson, D. Constructing validity: Basic issues in objective scale development. *Psychological assessment 7*, 3 (1995), 309–319.
9. Cortes, C., and Vapnik, V. Support-vector networks. *Machine Learning 20*, 3 (1995), 273–297.
10. Day, D., Gan, B., Gendall, P., and Esslemont, D. Predicting purchase behaviour. *Marketing Bulletin 2*, 5 (1991), 18–30.
11. De Choudhury, M.and Gamon, M., Counts, S., and Horvitz, E. Predicting depression via social media. In *ICWSM* (2013).
12. de Marneffe, M.-C., MacCartney, B., and Manning, C. D. Generating typed dependency parses from phrase structure parses. In *LREC* (2006).
13. Deci, E. L., and Ryan, R. M. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry 11*, 4 (2000), 227–268.
14. Diakopoulos, N. A., and Shamma, D. A. Characterizing debate performance via aggregated Twitter sentiment. In *Proceedings of CHI* (2010), 1195–1198.
15. Ford, J. K. *Brands laid bare: Using market research for evidence-based brand management*. John Wiley & Sons, 2005.
16. Friedman, J., Hastie, T., and Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software 33*, 1 (2010), 1–22.
17. Golbeck, J., Robles, C., Edmondson, M., and Turner, K. Predicting personality from Twitter. In *IEEE SocialCom* (2011), 149–156.
18. Golbeck, J., Robles, C., and Turner, K. Predicting personality with social media. In *CHI* (2011), 253–262.
19. Goldberg, L. R. A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe 7* (1999), 7–28.
20. Hauser, J. R., and Koppelman, F. S. Alternative perceptual mapping techniques: Relative accuracy and usefulness. *Journal of Marketing Research* (1979), 495–506.
21. IBM. InfoSphere BigInsights Social Data Accelerator, 2013.
22. Kohli, C. S., and Leuthesser, L. Product positioning: a comparison of perceptual mapping techniques. *Journal of Product & Brand Management 2*, 4 (1993), 10–19.
23. Kotler, P., and Armstrong, G. *Principles of Marketing*. Pretice Hall, 2013.
24. Ling, C. X., and Li, C. Data mining for direct marketing: Problems and solutions. In *KDD* (1998), 73–79.
25. Maslow, A. H. A theory of human motivation. *Psychological review 50*, 4 (1943), 370–396.
26. Maslow, A. H. *Personality and motivation*. Harlow, England: Longman, 1954.
27. Miller, G. A. Wordnet: A lexical database for english. *Communications of the ACM 38*, 11 (1995), 39–41.
28. Mukherjee, A., and Liu, B. Improving gender classification of blog authors. In *EMNLP* (2010), 207–217.
29. NetBase. Purchase Intent, 2013.
30. Nguyen, T., Phung, D. Q., Adams, B., and Venkatesh, S. Towards discovery of influence and personality traits through social link prediction. In *ICWSM* (2011).
31. Pennebaker, J. W., Francis, M. E., and Booth, R. J. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*.
32. Picard, R. W. Affective computing for HCI. In *HCI (1)* (1999), 829–833.
33. Quercia, D., Kosinski, M., Stillwell, D., and Crowcroft, J. Our Twitter profiles, our selves: Predicting personality with twitter. In *IEEE SocialCom* (2011), 180–185.
34. Rao, D., Yarowsky, D., Shreevats, A., and Gupta, M. Classifying latent user attributes in Twitter. In *SMUC* (2010), 37–44.
35. Ryan, R. M., and Deci, E. L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist 55*, 1 (2000), 68–78.
36. Ryff, C. D., and Keyes, C. L. M. The structure of psychological well-being revisited. *Journal of personality and social psychology 69*, 4 (1995), 719–727.

37. Sakaki, T., Okazaki, M., and Matsuo, Y. Earthquake shakes Twitter users: real-time event detection by social sensors. In *WWW* (2010), 851–860.

38. Schler, J., Koppel, M., Argamon, S., and Pennebaker, J. Effects of age and gender on blogging. In *CAAW* (2006).

39. Schwartz, S. H. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in experimental social psychology 25*, 1 (1992), 1–65.

40. Shi, L., Agarwal, N., Agrawal, A., Garg, R., and Spoelstra, J. Predicting US primary elections with Twitter. In *social* (2012).

41. Sismeiro, C., and Bucklin, R. E. Modeling purchase behavior at an e-commerce web site: a task-completion approach. *Journal of Marketing Research* (2004), 306–323.

42. Tay, L., and Diener, E. Needs and subjective well-being around the world. *Journal of Personality and Social Psychology 101*, 2 (2011), 354–365.

43. Tumasjan, A., Sprenger, T. O., Sandner, P. G., and Welpe, I. M. Predicting elections with twitter: What 140 characters reveal about political sentiment. In *ICWSM* (2010).

44. Venkatesh, V., and Agarwal, R. Turning visitors into customers: a usability-centric perspective on purchase behavior in electronic channels. *Management Science 52*, 3 (2006), 367–382.

45. Wahba, M. A., and Bridwell, L. G. Maslow reconsidered: A review of research on the need hierarchy theory. *Organizational behavior and human performance 15*, 2 (1976), 212–240.

46. Yarkoni, T. Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of research in personality 44*, 3 (2010), 363–373.

47. Zhang, Y., and Pennacchiotti, M. Predicting purchase behaviors from social media. In *WWW* (2013), 1521–1532.

Appendices

A. DEFINITIONS OF UNIVERSAL NEEDS

The following definitions are directly from or paraphrased from [15].

Structure: This need is about routine, consistency, systematicity, and predictability. It is "the need for things to be well organized and under control".

Practicality: "This need is about getting the job done, skill and efficiency. It also covers physical expression and experience, a celebration of the senses and the physical world."

Challenge: This need includes: (1) the urge for a personal sense of achievement by voluntarily taking up a challenge; (2) the urge for victory over an opponent; and (3) the urge for "authority and status or prestige".

Self-expression: This need is about expressing one's "ego, strength and independence" through external actions.

Excitement: "This need is about getting out there and living life, upbeat emotions, high energy, having fun."

Curiosity: This need is about the desire to discover and learn, "curiosity for its own sake".

Liberty: This need is about the thirst to have control over one's own actions without constraints, "a sense of total possibility and new things".

Ideals: This need includes: (1) the desire for refinement of style: "glamour and seduction, sophistication and elitism"; (2) the desire for refinement of principle, "a sense of community and responsible progress"; (3) the desire for refined spiritual development.

Harmony: This need is about appreciating and respecting "other people, their viewpoints and their feelings".

Love: "This need is all about social contact, whether one to one or one to many".

Closeness: This need is about the desire to nurture and to be nurtured, the sense of belonging.

Stability: This need is about physical security and control, reliability. "It has the sense of something sensible, tried and tested, with a good track record and a known history."

B. SAMPLE SURVEY ITEMS

For the need of **ideal**:

- *I have a yearning for glamour and sophistication.*
- *I crave refinement in life.*
- *I deeply appreciate anything that strives to reach the seemingly unattainable ideals.*
- *I tend to pursue and worship perfection.*

For the need of **closeness**:

- *I feel strong emotional needs for nurturing and being nurtured.*
- *I desires things that evoke warm memories and nostalgia.*
- *I want to create a caring environment for things that are close to my heart.*
- *I want things that strengthen the emotional tie with my loved ones.*
- *I enjoy all kinds of home-oriented activities because they make me feel close to my family.*