A Picture is Worth a Thousand Words: Improving Usability and Robustness of Online Recommendation Systems

Tiffany Hyun-Jin Kim   Virgil Gligor   Adrian Perrig

Carnegie Mellon University
{hyunjin1, virgil, adrian}@ece.cmu.edu

***Abstract*— Recent statistics show that the number of online shoppers are increasing where the majority of them use online recommendation systems for product/service reviews. Although online reviews are becoming increasingly important, consumers face two major challenges of usability and robustness when they make purchase decisions based on the available reviews. More specifically, usability issues arise when consumers need to be able to extract *relevant* information given a *high volume of data* with *uncertainty due to high variance*. For robustness, judging the degree of truthfulness of the available recommendations can be a daunting task for consumers. In this paper, we propose a post-purchase tracking system as an enhancement to current online recommendation systems by embracing a peer review process and ask each consumer to score the reviews that previous consumers have posted. Furthermore, we propose to visualize the peer review processes such that people find the recommendation systems more efficient and useful to learn information. Our preliminary user study results indicate that our post-purchase tracking system is a promising approach that can help online consumers determine what information to trust with confidence.**

***Keywords*– security, usability, online recommendation systems, visualization**

I. Introduction

According to the recently released Q2 2010 report on Online Shopper Intelligence, 83% of all consumers shop online at least once a week.[1] Compared to Q3 2009, the rate of online shoppers has increased by 31%. As online shopping becomes popular, online product reviews are more important than ever. According to an e-tailing group and PowerReviews study in May 2010, 50% of online shoppers conduct research online for at least half of their purchases and 64% consistently read online review before making purchase decisions.[2] Despite the increase in popularity and importance, two major challenges need to be addressed in online recommendation systems: usability and robustness.

Usability. One issue with online recommendation systems is the *high volume of data* and the way to display it. For some popular products, there can easily be hundreds, if not thousands, of reviews that other consumers have uploaded. As a result, given an extensive list of reviews, a prospective consumer would need to dedicate some time and effort to (at least) spot-read the reviews.

Another issue is related to the *relevance*: given possibly overwhelming amount of reviews, a consumer faces a challenge to determine the relevance of each review. Since each consumer has a unique personal taste, s/he needs to extract the most relevant information. Furthermore, it is possible that the product/service provider switches to a different product/service while the available reviews are still relevant to the previous product/service. Hence, the consumer would need to learn how relevant the available reviews are.

The third issue is the *uncertainty due to high variance*. An overwhelming amount of data can often be counter-productive, as the data may be overwhelmining in volume and self-contradictory. For example, reviews using the 5-star rating system can range from 1 star (lowest) to 5 stars (highest) based on the reviewers' personal opinion, and the 5-starred comments often portray contradictory information compared to the 1-starred comments. On the other hand, there are also products with few reviews, in which case consumers may not be able to find any useful information.

Robustness. Consumers make purchase decisions based on how much they are willing to try a product or a service after reviewing others' opinions. However, judging the level of truthfulness of the available reviews is an intricate task since there are various factors that affect the reviews. For example, critical people would provide negative reviews focusing on the downsides of a product/service, and some consumers with malicious intention may report false information.

Our approach. As a consumer, a challenge is to determine the relevance of each review. An observation is that recent purchasers who used the product/service can provide feedback on others' reviews so that other consumers can easily find more relevant and accurate information about the product/service. Furthermore, if such a reviewing system can be graphically represented with visual diagrams instead of a long list of reviews that current recommendation systems provide, people may find it more efficient and useful to learn information.

In this paper, we propose an initial approach to improve current online product review systems. Our post-purchase tracking system provides robust and accurate online product/service reviews by applying a peer-review process to recommendation systems as follows: recent product buyers provide feedback not only on the purchased product and the seller but also on the reviews that previous buyers have posted. Furthermore, we study ways to effectively visualize the tracking system such that consumers find it intuitive and helpful to determine what information to trust with high confidence.

This research was supported by CyLab at Carnegie Mellon under grants DAAD19-02-1-0389 and W911NF-09-1-0273 from the Army Research Office, and by support from NSF under the TRUST STC award CCF-0424422. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of ARO, CMU, NSF or the U.S. Government or any of its agencies.

[1]http://payment-times.com/wordpress/?p=765

[2]http://www.marketingprofs.com/charts/2010/3563/

Contributions. This paper presents our initial attempt to explore a new approach to secure current online recommendation systems by applying crowdsourcing such that users review and score other users' posted comments. In this manner, highly relevant information that the majority of consumers agree upon will be automatically emphasized and misleading information will be automatically corrected by consumers. We also explore how to effectively deliver the reviewed comments using graphical visualization.

II. Problem Definition

In this section, we delineate a precise problem definition, describe the desired properties for designing a post-purchase tracking system, and discuss our adversary model.

A. Problem Definition

We study how to improve current online recommendation systems for product or service reviews. Our fundamental question is: how can we design a robust online recommendation system that consumers can trust the presented information with high confidence? We also study approaches to visualize the enhanced recommendation system such that consumers find it intuitive and easy to learn the desired information.

B. Desired Properties

Robustness proportional to popularity. The enhanced recommendation system must be robust such that tampering with the system must be challenging. Also, the reviews must be non-inflatable such that adversaries cannot alter the reviews to convey benign consumers with misleading reviews. Moreover, it is desirable that the difficulty for tampering scales with product popularity such that the more popular a product is, the more challenging it is to tamper the review of the product.

Usability. The enhanced recommendation system must be intuitive for ordinary online shoppers to interpret and learn the desired information without difficulties. More specifically, the system should minimize the number of explicit interactions that a consumer needs to perform while comprehending the previous reviews. Also, the system should not require shoppers to possess special knowledge such that an ordinary shopper can easily interpret the presented information.

C. Adversary Model

Our paper addresses ways to appropriately capture correct information about a product/service such that online shoppers can make better decisions with high confidence. Based on this scope, we address an external adversary whose intention is to tamper with the current reviews such that benign shoppers perceive a fraudulent image of a product/service.

To tamper with the reviews, an adversary may attempt to emphasize misleading information on the review board. With heavily emphasized information, the adversary may succeed in deceiving a benign online shopper to change her decision in purchasing the product/service.

We consider two concrete attacks:

- **Sybil Attack.** An adversary may create multiple online virtual identities and inflate/deflate the reviews with emphasis on what he desires to deliver to benign shoppers [1].
- **Colluding Attack.** To convince a benign shopper with misleading information, two or more adversaries may collaborate to inflate each other's reviews or deflate others' reviews. With the representation of convincing opinions, the adversaries can successfully acquire the trust of the shopper.

III. Post-Purchase Tracking System

We introduce a post-purchase tracking system as an enhancement to the current online recommendation systems. Unlike the conventional recommendation systems, our system efficiently utilizes crowdsourcing to highlight relevant and trustworthy information.

A. System Overview

Our post-purchase tracking system is based on the peer-review process where community members review each other's comments and perform the following actions:

- Correct misleading information,
- Highlight valid information, and/or
- Add new information to enhance the current comments.

Based on this reviewing process, the information with the majority votes, which correctly represents the valid information, will be automatically accentuated; on the other hand, fallacious information will be abandoned and replaced with the correct information.

When a member is interested in a product/service that other members have peer-reviewed, she can gain trust in the product/service with high confidence based on the feedback that other members have contributed.

B. Example

Here is how our system protocol can be applied to an online recommendation system. When online shopper Alice purchases a digital camera, our post-purchase tracking system keeps track of an appropriate amount of time for Alice to try out the product. Later, our system invites Alice to participate in a brief survey which asks her to select the previously posted review(s) on the camera she read, if any, before she made her purchase decision.[3] Based on the selected reviews, the system further asks Alice to vote as follows:

- Each word in a review has two buttons: one for "agree" and one for "disagree."
- If Alice agrees with any word(s) in a review, she clicks the "agree" button.
- If Alice disagrees with any word(s) in a review, she clicks the "disagree" button.[4]

The system also asks Alice to leave further reviews that she did not find in the previously posted reviews.

Later whenever Bob is interested in purchasing the camera that Alice bought, Bob finds a collection of reviews that previous consumers posted to share their experiences with the seller and the camera. Rather than seeing an ordinary

[3]The invitations for review submission, where one can only submit a review for a product/seller after purchasing the product, is adapted from the current online recommendation systems such as eBay's.

[4]One could consider voting on keywords. However, a challenge is to automatically figure out the keywords that are important in the given context. While it may seem cumbersome that people rank each word, they would tend to pick the important words which will result in selecting the keywords.

review board that shows reviews one after the other, Bob now sees a review board with highlighted words that the previous consumers have voted to be highly relevant to this seller and the product. He also sees some de-emphasized words as a result of conflicts among previous reviewers. Thanks to the post-purchase tracking system where other reviewers contributed to strengthen and portray the correct information, Bob gets a clearer view about the product and gains trust in it with high confidence.

C. Protocol Description

In this section, we describe the post-purchase tracking system in detail. We define a review about a product to be a collection of words that a consumer posts such that others can refer to it and make better purchase decision. More specifically, let X_p be a set of all the reviewers who posted a review on product p in the system, and let a review by a consumer x on product p (i.e., $x \in X_p$) be the set of reviewer x's words $R_{(x,p)} = \{W_i^{(x,p)}\}$ where $W_i^{(x,p)}$ is the score of i^{th} word in consumer x's review $R_{(x,p)}$ on product p. When x creates review $R_{(x,p)}, W_i^{(x,p)} = 1 \ \forall i \leq |R_{(x,p)}|$ (i.e., each word gets a score of 1).

When consumer y purchases product p, the system sends a reminder to y to participate in the post-purchase survey to provide feedback on not only the product p but also the reviews $R_{(x,p)}$ that were posted by other consumers x (where $x \neq y$) that y referred to. More formally, let $i^y_{(x,p)} = 1$ if y read x's review on product p, 0 otherwise. Then, $S^y_{R_{(x,p)}} = \{R_{(x,p)} : x \in X_p \wedge i^y_{(x,p)} = 1\}$ (i.e., a set of the product p's reviews that y referred to before purchasing p). Now for all the reviews that y referred to, the system requests y to vote on the word(s) that y agrees with and update the score(s) on the word(s). In other words, $\forall R_{(x,p)} \in S^y_{R_{(x,p)}} \forall i \leq |R_{(x,p)}|$,

$$W_i^{(x,p)} = \begin{cases} W_i^{(x,p)} + 1 & \text{if } y \text{ agrees with } W_i^{(x,p)} \\ W_i^{(x,p)} - 1 & \text{if } y \text{ disagrees with } W_i^{(x,p)} \end{cases}$$

If a significant number of people agree with a word on a review (i.e., $W_i^{(x,p)}$ on $R_{(x,p)} \geq tt_{up}$ where tt_{up} is an upper bounding threshold), then the post-purchase tracking system emphasizes the word such that the succeeding consumers pay closer attention to the emphasized word. Similarly, if a significant number of people disagree with a word on a review (i.e., $W_i^{(x,p)}$ on $R_{(x,p)} < tt_{low}$ where tt_{low} is a lower bounding threshold), then the post-purchase tracking system de-emphasizes the word such that the succeeding consumers pay minimal attention to the disagreed word. In case y wants to make a new contribution and provide feedback that has not been mentioned by others, our system allows y to leave further comments. In the next section, we delineate how to visualize such emphasis on words.

IV. Recommendation Visualization

We now explore visual approaches to effectively deliver the accurate information that our post-purchase tracking system extracts from the crowdsourcing process. We suggest two initial approaches to visualize the system: typographic emphasis and grouping.

Figure 1. An example of a modified Amazon.com's product review page using the post-purchase tracking system with typographic emphasis for an effective visualization. In this example, key words in bold red italics (e.g., delay, low price, cheap) are what reviewers have agreed with and the size of the words is proportional to the degree of agreement. On the other hand, words with lighter color and with a strikethrough line (e.g., blurry pics, sturdy) represent what reviewers have disagreed with and the size of the words is inversely proportional to the degree of disagreement. A word that people do not uniformly agree/disagree (i.e., very quick) is highlighted in green italics with a question mark.

A. Typographic Emphasis Approach

One method to differentiate words is by exaggerating them in different fonts from the rest of the text. Such typographic emphasis is effective as human eyes are receptive to differences in brightness.[5] Based on the brightness, size, color, and/or fonts of the text, an emphasis can be placed on the text. By using typographic emphasis on our post-purchase tracking system, the words that consumers agreed with can be emphasized with bolder, bigger, and vivid-colored fonts whereas the words that consumers disagree with can be de-emphasized with lighter, smaller, and dull-colored fonts. Figure 1 is a modification of an actual Amazon.com's product review page where the degree of brightness, size, color, and fonts is correlated with the number of votes that reviewers agree/disagree. Words that are conflicting among reviewers are represented in green italic fonts with a question mark. Moreover, the emphasized/de-emphasized words can carry extra information such as the percentage of the reviewers who agreed/disagreed with the words, for example by placing a mouse pointer over the words.

B. Grouping Approach

Unlike the typographic emphasis approach that exaggerates agreed/disagreed words in different style from the rest of the text, this approach groups those words that reviewers have edited in three categories: "strongly agreed", "strongly disagreed", and "conflicting". In the "strongly agreed" category, the words that the majority of reviewers have agreed with and hence having high $W_i^{(x,p)}$ scores are displayed in dark, big, and vivid-colored fonts. In the "strongly disagreed" group, the words that the majority of reviewers have disagreed with and hence having low $W_i^{(x,p)}$ scores are displayed in light, small, ands dull-colored fonts. Those words that some

[5]http://vudat.msu.edu/teach/page-design

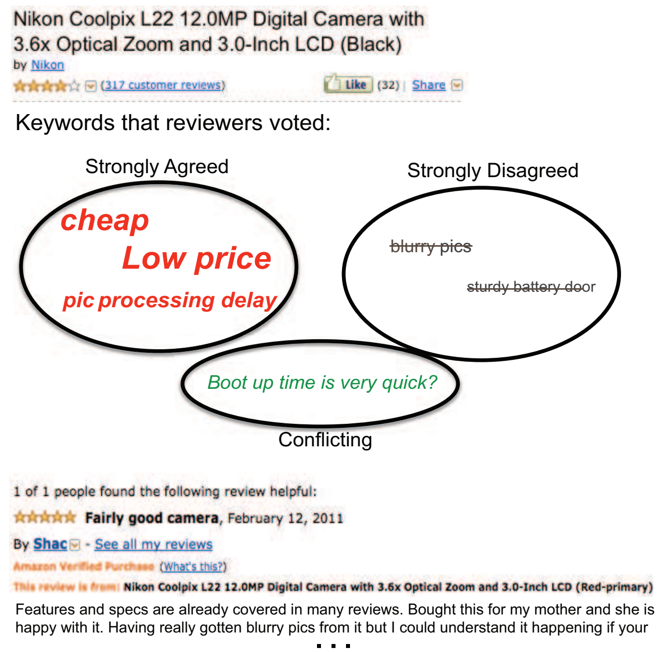

Figure 2. An example of the same review as Figure 1 but using the grouping approach. In this example, the words that reviewers edited are grouped into three categories: strongly agreed, strongly disagreed, and conflicting. If consumers are further interested in how these words are placed, they can place a mouse over each word to learn the actual percentage of reviewers who voted on it. Furthermore, if consumers are interested in the opinion in the context, they can click the word (or phrase) to find a list of all the actual reviews that contain it. The grouping diagram is followed by the actual reviews.

reviewers consider to portray the product/service correctly but other reviewers consider to be misleading, hence having high variance $W_i^{(x,p)}$ scores are placed in the "conflicting" group. The diagram of the three groups are inserted before the actual reviews. Figure 2 is an example of a product review using the grouping approach. Similar to the typographic emphasis approach, the words in each group can carry extra information such as the percentage of the reviewers who agreed/disagreed with the words, for example by placing a mouse pointer over the words. Furthermore, if consumers are further interested in getting the actual context for a particular word within a group, clicking the word will result in a list of all the reviews that actually contain the word.

C. Discussion

Our two initial visual approaches for the post-purchase tracking system are intuitive to interpret and learn the desired information without difficulties: 1) with words that are emphasized in bigger, bolder, and vivid-colored fonts, ordinary consumers can easily acquire the information that peer-reviewers have mostly agreed with, 2) with words that are de-emphasized with smaller, dull-colored strikethrough fonts, consumers can easily learn that these words are likely to be irrelevant with the products, and 3) with words that are in italics followed by question marks, consumers can easily guess that these words are controversial and hence they should not heavily depend on these words. In particular, while our typographic emphasis approach delivers these three points within the context, our grouping approach places heavy emphasis only on what other peer-reviewers have edited. As a result, our visual approaches can effectively deliver the merits of the post-purchase tracking system and consumers may find it easy to interpret.

V. Security Analysis

We analyze how resilient our post-tracking system is against the active attacker model as described in Section II-C. We also analyze whether our system satisfies the desired properties as mentioned in Section II-B. Although it is not necessarily an adversary model, edits by multiple reviewers may not uniformly agree on certain words, in which case fluctuation in the scores on the words occurs. We analyze how our system mitigates such high variance in some conflicting words.

A. Analysis on Adversary Model

Sybil attack. An adversary can launch a Sybil attack by creating multiple online virtual identities to inflate/deflate reviews such that benign shoppers are exposed to misleading information. However, our post-purchase tracking system invites a consumer who has indeed purchased a product to participate in the peer-reviewing survey.[3] As a result, unless an adversary purchases the product for each of his virtual identity, the adversary cannot inflate/deflate reviews using multiple online virtual identities.

Colluding attack. Two or more adversaries can collude to inflate each other's reviews or deflate others' comments, and successfully emphasize misleading information about a product. However, unless the majority of the reviewers collude, our post-purchase tracking system is resilient against colluders since their inflation/deflation is proportional to the number of colluders and hence limited. Moreover, other legitimate reviewers are highly likely to disagree with the colluders, in which case their scores may lead to either contradiction with the colluders' or over-rule the colluders' inflated/deflated opinion.

Inflating/downgrading attack. An adversary may attempt to inflate or downgrade a product by creating controversy with the most positive/negative comments. However, unless enough number of people (i.e., about same as the number of people who agree with the comments) have the same objective to inflate/downgrade the product, our system would not categorize the comments as conflicting. Hence, the adversary's action will be buried by others' votes for the comments with high probability.

B. Analysis on Desired Properties

Robustness proportional to popularity. As Section V-A shows, our system is robust against active attackers who attempt to tamper with the system and the degree of inflating/deflating the reviews is limited, at most to the number of colluding attackers. Moreover, people enjoy providing feedback and prefer to take some action for the community. For example, a study shows that people actively report phishing websites even though a well-known search engine Google already acknowledges them [2]. As people feel good when they can perform positive actions for the common good, they may enjoy providing feedback on recently purchased products for future consumers who are considering the same products. Since people enjoy providing feedback, mostly voluntarily but sometimes with small incentives, consumers who read the feedback can trust the information with high confidence.

An interesting case is when consumers indeed disagree with the prior posts that they refer to before they purchase a product. In this case, people are more likely to act since they have a vested interest to rectify the injustice that was inflicted onto them. As a consequence, the system exhibits the properties of altruistic punishment where people have incentives to behave in a manner that benefits the society overall. A prior work also shows that users had an incentive to complain [2]. Therefore, people are more likely to act and correct misleading information, securing the system in a proportional manner to the popularity of the product.

Usability. Our post-purchase tracking system revolves around a similar process as the current review system by inviting product purchasers to participate in the follow-up surveys and provide feedback. The major difference is that our system requests consumers to identify the reviews that they read before their purchases and score those reviews to strengthen the merits that the products provide and to correct misleading information given "agree" and "disagree" buttons in the reviews. Hence, our system is intuitive since the consumers are asked to perform the following actions: they are asked to click "agree" if their experience on the products is agreeable with the particular word/phrase in the review text; on the other hand, the consumers are asked to click "disagree" if their experience contradicts with the word/phrase in the review text.

In terms of the visual representation of the post-purchase tracking system, our two initial visual approaches are intuitive to interpret and learn the desired information with varying brightness, size, color, and fonts of the edited words. Hence, consumers may find it trivial to learn the information that the system presents.

C. Analysis of High Variance Scores

Some words in a review text may cause conflicts among peer reviewers; some may agree with the words but some may have different opinion. In order to address such cases, our system has a third category called "conflicting" where words are followed by a question mark to indicate that these words are in conflict among reviewers, causing high variance. Hence, consumers are advised to consider the controversial words in mind before they make purchase decisions.

VI. Evaluation

As an initial attempt to evaluate our approaches, we conducted a small user study. In this study, our objectives are 1) to analyze whether consumers gain trust with high confidence in the reviews using the post-purchase tracking system, 2) to analyze the usability of our system, and 3) to get feedback on how to improve our initial approaches.

Demographics and background information. We conducted an interview with 20 participants, 10 males and 10 females within the age range of 19–56 who are either university staffs or students with a college degree or higher. They were all active online product buyers where 10% of the participants purchase product on a daily basis, 15% purchase products weekly, 60% purchase products monthly, and 15% purchase products once every three months. Everyone responded that they would research about the product, for example using search engines, before they would make the purchase decision. Also, everyone responded that they read the product reviews to find useful information that the product description would not mention about. Given a likert scale from 1 (not at all) to 5 (very much), people somewhat rely on the product reviews that others post (3.8 out of 5 on average). We also asked a question about how much they trust the information on the product reviews and from the same 5-point likert scale, the average was 3.5.

Procedure. To verify whether our system with the visual approaches enhances the robustness and the usability, we conducted a comparison study with the current Amazon.com's review system. More specifically, we prepared three interfaces: one with Amazon.com's review style, one with our system's review mechanism with the typographic emphasis approach (e.g., Figure 1), and one with our system's review mechanism with the grouping approach (e.g., Figure 2). Since electronics are one of the best selling items on Amazon.com,[6] we picked a camera as a product for the interview. We began the interview using an existing Amazon.com's camera review as a baseline and uniformly randomized the order of the other two interfaces to minimize the possibility of building biases on the interfaces, such as the tendency of favoring the interface that is presented at last. After explaining each interface, we asked some questions and we requested the participants to speak outloud during the interview session. Below are sample questions that we asked for each interface:

- What do you like/dislike about this review style?
- What do you recommend to improve this review style?
- How easy is it to read this review? (5-pt scale)
- How useful do you find this review? (5-pt scale)
- Is this review style intuitive and easy to understand? (5-pt scale)
- Is this review style helpful to trust the information about this product with high confidence? (5-pt scale)
- Among the three review styles, which one do you prefer? Why?

Results. Majority of the participants (85%) provided positive responses and feedback on our approaches. For the baseline Amazon.com's review prototype, 3 people expressed their preference of reading the story and 1 person liked the capability of entering freetext rather than filling in a formatted survey. However, other 17 people (85%) raised the issue of reading lengthy text, which takes significant amount of effort to gather the information they want.

For the review prototype using the typographic emphasis visualization, people liked the highlighted keywords that grabbed their attention immediately while they could still read the context to get the detailed information if they wanted. They also expressed interest in the crowd agreement/disagreement approach to raise the quality of the reviews. However, 3 people raised concerns that the changes in font sizes and colors distracted them from reading the reviews and 1 person suggested to use only bold fonts for emphasis without changing the font size.

[6] http://www.amazon.com/gp/bestsellers

Table I

INTERVIEW RESULTS OF THREE REVIEW STYLES. EACH VALUE REPRESENTS THE AVERAGE OF THE 20 PARTICIPANTS' RESPONSES FROM A 5-POINT LIKERT SCALE.

	Amazon.com	Typo-emphasis	Grouping
Ease of reading	4	3.7	4
Usefulness	3.8	3.8	3.8
Intuitiveness	3.8	4	4
Helpfulness to trust	3	3.8	3

For the prototype using the grouping visualization approach, people enjoyed the simplicity and the comparison feature. They also stressed the advantage of being able to learn the general summary about the pros and cons of the product fast rather than reading sometimes lengthy reviews; 1 participant mentioned that this style was user-friendly. However, 3 people raised concerns about having to learn this new review style and they suggested placing a legend which explains what each group means and how they can get further information, such as by clicking the terms, etc. Also, 2 people raised the usability concern about the extra clicking on each term to gather more information.

The result for the sample questions are summarized in Table I. This table shows that both of our visual approaches are more intuitive and easier to understand than the current Amazon.com's review approach. Furthermore, people expressed greater helpfulness to trust the information about the product with high confidence using our system with the typographic emphasis visualization.

Discussion. Based on this preliminary interview results, our approaches are at least as useful and as helpful to make people trust the information about a product with high confidence as the current Amazon.com's review approach (refer to Table I). Moreover, out approaches are more intuitive and easier to understand than that of Amazon.com.

Only 3 out of 20 people preferred the current Amazon.com's review style, merely because they were concerned that others would not easily interpret the visual representations. However, they mentioned that a legend explaining the visual effects would make our approaches be more understandable. Also, some people suggested combining two of our approaches such that the grouping visual approach is supported by the typographic emphasis approach when users click the terms for further details.

VII. RELATED WORK

Recommendation systems apply collaborative filtering techniques to enable the prediction of user preferences [3]–[5]. Most online recommendation systems can be classified as user-based collaborative filtering system, where a social network is created among those who share the same rating pattern, and the recommendation is provided to the user based on the item rated by the most similar user [3], [5], or item-based collaborative filtering system, where the prediction is computed based on the similarly rated items by a target user [6]. One well-known example of the item-based collaborative system is the Amazon.com Recommendations, which incorporates a matrix of the item similarity [7]. Other technologies, such as nearest neighbor methods [3], [8], have also been applied to recommender systems. Unlike these recommendation systems, our approach relies on peer-review process where each reviewer contributes to score posted reviews such that representative information is emphasized while misleading information is corrected.

Many researchers have worked on enhancing trustworthiness of the online recommendation systems [9]–[12]. For example, undirected transaction graphs [9], [10] and the reputation network constructed from buyers' feedback [11] are used to identify fraudulent users. Chiou et al. uses social networks to provide authentic online reviews [12]. Rather than depending on the strongly-tied social networks of the consumers and decide trustworthiness of the reviews, our system depends on weak-tied crowdsourcing for reviews that consumers can trust with high confidence.

VIII. CONCLUSION

In this paper, we explore a new approach to enhance current online recommendation systems. We introduce the post-purchase tracking system where current recommendations are scored by peer reviewers to 1) emphasize relevant information, 2) correct misleading information, and 3) provide new information to enhance current reviews. To communicate the opinions of other community members, we introduce two initial visualization approaches. Based on the preliminary user study result of our post-purchase tracking system and two visual diagrams, our next step is to implement our system on an actual online recommendation system for larger-scale evaluation. We anticipate that our initial explorations in this area will encourage more research that will ultimately enable users to trust reviews with high confidence.

REFERENCES

[1] J. R. Douceur, "The Sybil Attack," in *IPTPS*, 2002.

[2] C. Kuo, "Reduction of End User Errors in the Design of Scalable, Secure Communication," Ph.D. dissertation, Carnegie Mellon University, 2008.

[3] P. Resnick, N. Iacovou, M. Suchak, and P. B. andJohn Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," in *Proceedings of CSCW*, 1994.

[4] W. Hill, L. Stead, M. Rosenstein, and G. Furnas, "Recommending And Evaluating Choices In A Virtual Community Of Use," in *Proceedings of CHI*, 1995.

[5] U. Shardanand and P. Maes, "Social Information Filtering: Algorithms for Automating "Word of Mouth"," in *Proceedings of CHI*, 1995.

[6] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based Collaborative Filtering Recommendation Algorithms," in *Proceedings of WWW*, 2010.

[7] G. Linden, B. Smith, and J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet Computing*, vol. 7, pp. 76–80, 2003.

[8] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," in *Proceedings of the 14th Conference on Uncerrainty in Artificial Intelligence*, 1998.

[9] B. Zhang, Y. Zhou, and C. Faloutsos, "Toward a Comprehensive Model in Internet Auction Fraud Detection," in *Proceedings of the 41st Hawaiian International Conference on System Sciences*, 2008.

[10] A. Pandit, D. H. Chau, S. Wang, and C. Faloutsos, "NetProbe: A Fast and Scalable System for Fraud Detection in Online Auction Networks," in *Proceedings of WWW*, 2007.

[11] Y.-H. Kuo, F. Wu, and R.-L. Hsiao, "A Trust Evaluation to Help On-line Consumer to Choose the Provider in Auction Web Site," in *Proceedings of the International Conference on Service Systems and Service Management*, 2007.

[12] X.-Y. Chiou, S.-Y. Chang, G. Mezzour, A. Perrig, and H.-M. Sun, "A Trustable Reputation Scheme Based on Private Relationships," in *Proceedings of ASONAM*, 2009.