Crowdsourcing Quality Control of Online Information: A Quality-Based Cascade Model

Wai-Tat Fu and Vera Liao

Department of Computer Science, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
wfu/liao28@illinois.edu

Abstract. We extend previous cascade models of social influence to investigate how the exchange of quality information among users may moderate cascade behavior, and the extent to which it may influence the effectiveness of collective user recommendations on quality control of information. We found that while cascades do sometimes occur, their effects depend critically on the accuracies of individual quality assessments of information contents. Contrary to predictions of cascade models of information flow, quality-based cascades tend to reinforce the propagation of individual quality assessments rather than being the primary sources that drive the assessments. We found even small increase in individual accuracies will significantly improve the overall effectiveness of crowdsourcing quality control. Implication to domains such as online health information Web sites or product reviews are discussed.

Keywords: Web quality, user reviews, information cascades, social dynamics.

1 Introduction

The rapid growth of the WWW has led to increasing concern for the lack of quality control of online information. Many have suggested that quality assessments of online information could be "crowdsourced" to the public by harnessing the collective intelligence of online users by allowing them recommend contents and share the recommendations with others. Although this kind of crowdsourcing seems to work well in many domains [2, 3, 9], recent studies, however, show that aggregate behavior is often subject to effects of cascades, in which the social dynamics involved during the accumulation of recommendations may not be effective in filtering out low-quality information [8]. Indeed, previous models show that when people imitate choices of others, "bad cascades" may sometimes make low-quality information more popular than high-quality ones [1].

Previous cascade models often assume that users imitate choices of others without direct communication of quality information. We extend previous models by assuming that users can make multiple choices among web sites, and they can communicate their quality assessments to other users. Our goal is to understand how individual quality assessments may moderate effects of information cascades. We investigate how cascades may be influenced by factors such as accuracies of individual quality assessments, confidence levels of accepting other users' recommendations, and impact of

J. Salerno et al. (Eds.): SBP 2011, LNCS 6589, pp. 147–154, 2011.

aggregate user recommendations on choices. The goal is to simulate effects at the individual level to understand how the social dynamics may lead to different aggregate patterns of behavior. Specifically, we investigate how individual quality assessments propagate in a participatory Web environment, and whether information cascades may moderate the effectiveness of aggregate user recommendations, such as whether they may lead to "lock in" to low-quality information; and if so, to what extent will they impact the overall effectiveness of crowdsourcing quality control of information.

2 Background

2.1 Quality Assessment of Online Information

While research has shown that people are in general good at choosing high quality Web sites, it is also found that people often browse passively by relying on what is readily available. For example, health information seekers are found to rely on search engines to choose online sources, and may utilize surface features (such as layout and structures of web pages) to infer quality [5]. Accuracies of quality assessments are also found to vary among individuals, and can be influenced by various factors such as background knowledge, Internet experience, cognitive resources available, etc [4]. Liao & Fu [7] also found that older adults, who tended to have declined cognitive abilities, were worse in assessment quality of online health information.

Research in the field of e-commerce shows that user reviews can often significantly influence choices of online information, and their impact may be biased by "group opinions" while ignoring their own preferences [3]. Research also shows that effects of cascades and social dynamics may influence effectiveness of user reviews [8] in guiding users to choose high-quality products. However, to our knowledge, none has studied how cascades may be influenced by accuracies of quality assessment at the individual level, which may lead to different emergent aggregate behavioral patterns. In fact, research has shown that in certain situations, even random accumulation of user recommendations may induce more users to follow the trend, creating a "rich gets richer" cascading effect that distorts the perception of quality. On the other hand, previous cascade models often do not allow users to pass their quality assessments to other users, and it is possible that the passing of quality information may moderate effects of cascades. The goal of the current model is to test the interactions between individual and aggregate quality assessments to understand the effectiveness of collective user recommendations on promoting high quality information.

2.2 Cascade Models

Many models of information cascades have been developed to understand collective behavior in difference contexts, such as information flow in a social network [6], choices of cultural products in a community [1], or effectiveness of viral marketing in a social network [10]. In the model by Bikhchandani et al [1], information cascades are modeled as imitative decision processes, in which users sequentially choose to accept or reject an option based on their own individual assessment and the actions of other users. Each user can then observe the choice made by the previous users and infer the individual assessments made by them. In most information cascade models, the users can only observe the choices made by previous users and learn nothing from the users about the quality of the options. Given this limited information, a string of bad choices can accumulate, such that a cascade might lock future users on a low-quality option. Similarly, cascade models of information flow aim at characterizing how likely information will spread across a social network [6, 10] without any specific control on how likely one may assess the quality of the information and how the passing of the quality information to other users may influence how it spreads to other people in the network. It is therefore still unclear how the passing of quality assessments by users, in addition to their choices, might influence information cascades, and how the extent to which the aggregate quality assessments can be relied on to filter out low-quality information.

3 The Model

The general structure of the model is to have a sequence of users each deciding whether to select one of the 20 sites (or select nothing). When user *i* selects a site *j*, a recommendation will be given to the site based on the user's imperfect quality assessment. User *i+1* will then interpret the recommendation by user *i* with a certain level of confidence to determine whether to select the same site *j* or not. If user i+1 does not select j, she will either select a site from a recommended list of sites based on aggregate recommendation by previous users or select nothing. The model therefore aims at capturing the interaction of "word of mouth" information from the neighboring users and the global assessment of quality by previous users.

3.1 The Web Environment

We simulated the environment by creating sets of 20 Web sites, with each set having different proportions of sites with good and bad quality. A user can freely pick any site (or not pick any) and give a positive or negative recommendation (or no recommendation) to it. This recommendation is available for the next user, acting as a local signal (i.e., word-of-mouth information) collected from neighboring users. Recommendations will also be accumulated and used to update the list of most recommended sites, which will be available for the next users. The choice of sites and recommendation are therefore influenced by three factors: (1) the perceived quality of the site by the user (the private signal), (2) the recommendation provided by the previous user (the local signal), and (3) the accumulated list of most recommended sites (the global signal).

3.2 A Quality-Based Cascade Model

We simulated the sequential choice and recommendation process by a set of 1000 users as they chose and gave recommendation to the set of 20 Web sites. When the simulation began, a user was randomly assigned to a site. The user then evaluated the site, and based on the user's perceived quality, it will assign a positive or negative recommendation to the site. Specifically, it was assumed that when the site had good quality, there was a probability Q (or $1\text{-}Q$) that the user would perceive that it had good (or bad) quality. When the perceived quality was good (or bad), there was a probability *R* (or *1-R*) that the user would give a positive (or negative) recommendation. The next user could then interpret the recommendation given by the previous user to infer the quality of the site. If the recommendation was positive (or negative), there was a probability *C* (or *1-C*) that the next user would choose (or not choose) the same site. If the user did not choose the same site, there is a probability *L* (or *1-L*) that it would randomly select from the top recommended list (or not choose any site). Fig. 1 shows the decision tree of the model.

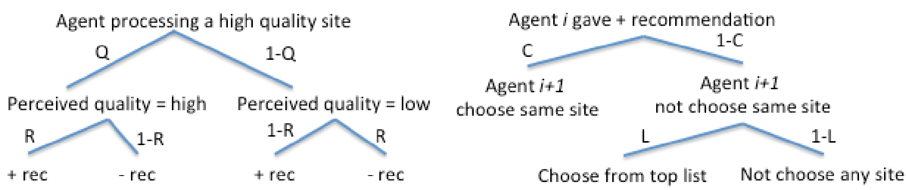

Fig. 1. The quality-based cascade model

3.3 Choice Accuracies

Fig. 2 shows the mean proportions of choice of high quality sites by the users with different values of Q (0.5, 0.55, 0.6) in a High (20% low), Medium (50% low), and Low (80% low) quality environments, with (a) low (C=0.5) and (b) high (C=0.9) confidence levels. In the "Top1" environment, when users decided to select from the recommendation list, she always picked the site with the highest recommendation; in the "Top10" environment, the user randomly picked a site among the best 10 sites.

Results show significant interactions among the private, local, and global signals. Good individual quality assessments (i.e., Q) in general lead to better overall accuracies in choosing high quality sites, but its effect on choice is strongly magnified by both the global signal (i.e., the recommendation list) and the confidence level in neighboring users' recommendation (i.e., C). Even when the confidence level is low (C=0.5), good quality assessments can be propagated through the aggregate recommendation list, assuming that other users will select the top recommended sites (in the Top1 environment). However, when users randomly selected sites from a longer list (Top10 environment), this channel of propagation diminished quickly, as shown by the low accuracies in the Top10 environment (accuracies in Top5 were somewhere between Top1 and Top10). When the confidence level was high (C=0.9), good quality assessments can be propagated through "word of mouth" local information, which improved the overall accuracies even in the top-10 environment. In general, even a small increase of Q from 0.5 (random choice) to 0.55 will lead to significant increase in overall choice of the correct Web sites for the aggregate. *Results demonstrate how local and global signals can magnify individual assessment of quality to increase the overall effectiveness of crowdsourcing quality control.*

3.4 Effects of Information Cascades

Fig. 3 shows examples of the simulations that illustrate whether information cascades occur in different conditions. Each figure shows the aggregate (sum of positive and negative) recommendations of the 20 web sites (different color lines) in the Medium

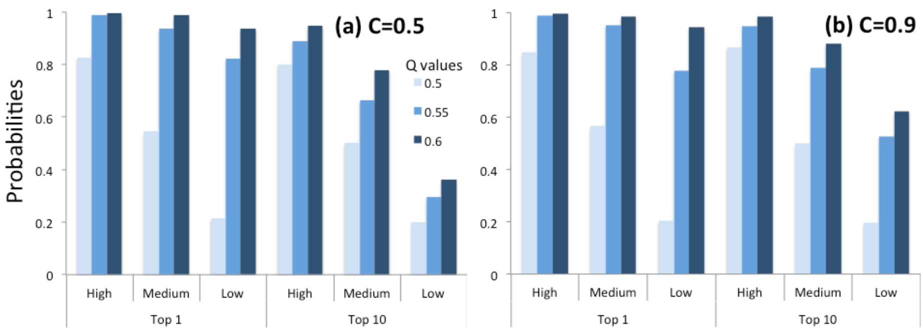

Fig. 2. Choice accuracies of high quality sites by users with different quality assessment accuracies (Q=0.5, 0.55, 0.6) and with low (a) and high (b) confidence levels (C) of other's recommendation, in environments with High (20%), Medium (50%) and Low (80%) ratio of low quality sites. Increasing Q from 0.5 to 0.55 leads to sharp increase in performance, especially in the Top1 environment and when C is high.

quality environment. Information cascades occur in all conditions, except when Q=0.5, C=0.5 in the Top10 environment, in which apparently there is no clear direction given by each user. In contrast, in the Top1 environment (similar results obtained for different values of C), information cascades do occur through the selection of the most recommended sites. When Q=0.5, because quality assessments are at chance level, information cascades occur whenever successive users choose and assign a positive recommendation on a random site and made the site the most recommended, which was enough to drive other users to keep selecting it. Higher Q value in Top1 environment shows even stronger information cascade, as consistent positive recommendations quickly lock on to a high quality site.

In the Top10 environment, when C=0.9, information cascade occur through the propagation of local "word-of-mouth" information. However, when Q=0.5, information cascade occurs much slower, as it requires successive random assessment of high quality of a particular site, but even when that happens, the effect is weak and tends to fade away quickly, as there is no reinforcement by the global signal. However, when Q is high, information cascades make the high quality sites quickly accumulate positive recommendation and increase their chance of being selected from the top-10 list. In other words, the pattern of results suggests that, when quality assessment is accurate, high quality sites tend to stand out quickly from the low-quality ones. Growth in positive recommendation for high quality sites quickly reinforces the flow of quality assessments by individuals. However, cascades themselves do not drive quality assessments to a sustainable level unless individual assessments of quality are low, and even so it requires a relatively rare sequence of random events to induce the cascades.

3.5 Effects of Initial Conditions

The above simulations assumed that all sites received no recommendation initially and users randomly choose among them to recommend to others. This may be different from actual situations as "word-of-mouth" information is often distributed

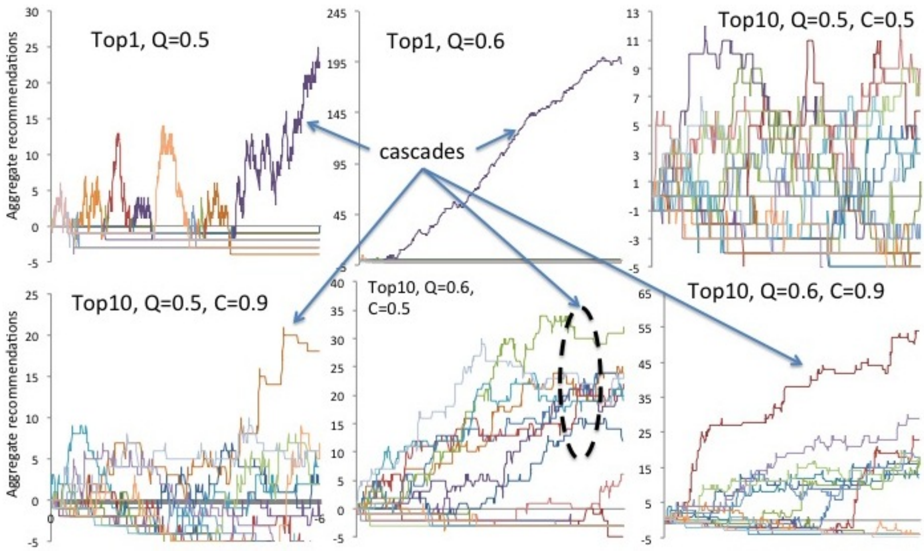

Fig. 3. Recommendations (y-axis) given to the sets of Web sites sequentially by 1000 users (x-axis) in the Medium environment. Color lines represent the net aggregate recommendations of 20 Web sites. Cascades occur when one (or more) Web site dominates others as more user recommendations are added. Except when Q=0.5, cascades tend to favor high-quality sites.

unevenly, or is subject to viral marketing as people are promoting their Web sites. We therefore simulated the effects of having positive recommendation on a subset of low quality sites and see how well these recommendations could hold up across time.

Fig. 4 shows the simulation results in the Medium quality environment with 2 of the low quality sites started off with 20 positive reviews (a higher number simply takes more cycles (users) to reach the same patterns). Consistent with previous simulations, when Q=0.5, the initially positively recommended low quality sites never come down, as users are not able to pass useful quality assessments to other users. In fact, bad cascades tend to drive up these low-quality sites. However, when Q is slightly higher (slightly more than random chance), recommendations of these sites all come down quickly as users collectively give negative recommendations on them. In the Top1 environment, sequential assignments of negative recommendations by multiple users quickly bring down the recommendations for the low quality sites and remove them from the top-1 list. When this happens, positive recommendations accumulate quickly on a high-quality site. In the Top10 environment, even when Q=0.5, when C=0.9, word-of-mouth information may accumulate and lead to cascades, although it is equally likely that the "winning" site has high or low quality. When Q is high, however, recommendations on low-quality sites come down quickly, while high-quality sites accumulate positive recommendations and become more likely to be selected once they appear on the top-10 list. This effect again was magnified by a higher C value, *showing that both local and global information significantly reinforces the collective quality assessments by multiple users.*

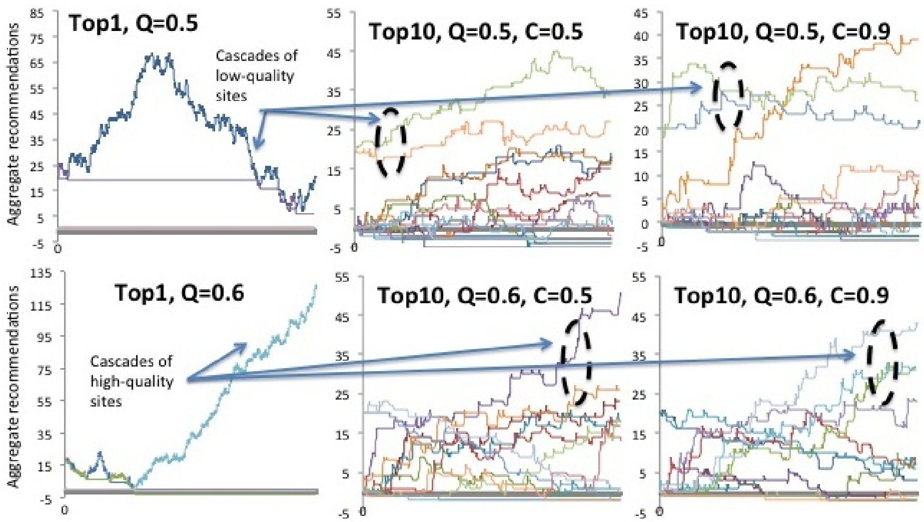

Fig. 4. Recommendations given to the sets of Web sites sequentially by 1000 users in the Medium quality environment. Two of the low-quality sites had 20 positive reviews when the simulation started. When Q=0.5, these low-quality sites tend to stay high, and “bad” cascades often drive them even higher; When Q=0.6, these low-quality sites quickly come down, and “good” cascades tend to magnify aggregate recommendations of high-quality sites.

4 Conclusions and Discussion

We extend previous cascade model to understand how quality assessments of information may be magnified by the effects of cascades at both the local (“word-of-mouth”) and global (aggregate recommendation list) levels. We found that in general, even when individual assessments of quality is only slightly better than chance (e.g., p=0.55 or 0.6 that one can correct judge quality), local and global signals can magnify the aggregated quality assessments and lead to good overall quality control of information, such that users can more likely find high quality information. We also show that when users can pass quality assessments to others, *cascades tend to reinforce these assessments, but seldom drive the assessments to either direction*. The results at least partially support the effectiveness of crowdsourcing quality control: even when users are far from perfect quality assessment, so long as the overall quality assessments can flow freely in the social environment, effects of the aggregated quality assessments can be magnified and are useful for improving the choice of high-quality Web sites.

The current model allows users to choose among many options (or choose nothing), which seems to make cascades less likely to occur (compared to a binary choice), but also more unpredictable. The mixing of information cascades of many scales among competing choices results in a turbulent flow to outcomes that cannot be easily predicted [8], but requires a model to predict the multiplicative effects among different signals. Our model shows that cascades are not necessarily bad, they could be effective forces that magnify the aggregated effort of quality assessments (even though individually they are far from perfect) to facilitate quality control of online information. Designs should therefore encourage more efficient flows of user recommendations to fully harness the potential of crowdsourcing of quality control.

References

1. Bikhchandani, S., Hirshleifer, D., Welch, I.: A theory of fads, fashion, custom, and cultural change in informational cascades. Journal of Political Economy 100(5), 992–1026 (1992)
2. Brossard, D., Lewenstein, B., Bonney, R.: Scientific knowledge and attitude change: The impact of a citizen science project. International Journal of Science Education 27(9), 1099–1121 (2005)
3. Cheung, M., Luo, C., Sia, C., et al.: Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. Int. J. Electron. Commerce 13(4), 9–38 (2009)
4. Czaja, S.J., Charness, N., Fisk, A.D., et al.: Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement (create). Psychology & Aging 21(2), 333–352 (2006)
5. Eysenbach, G.: Credibility of health information and digital media: New perspective and implications for youth. In: Metzger, M.J., Flanagin, A.J. (eds.) The John d. And Catherine t. Macarthur Foundation Series on Digital Media and Learning, pp. 123–154. MIT Press, Cambridge (2008)
6. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, D.C. (2003)
7. Liao, Q.V., Fu, W.-T.: Effects of cognitive aging on credibility assessments of online health information. In: Proceedings of the 28th Annual ACM Conference on Computer-human Interaction (CHI), Atlanta, GA (2010)
8. Salganik, M.J., Dodds, P.S., Watts, D.J.: Experimental study of inequality and unpredictability in an artificial cultural market. Science 311(5762), 854–856 (2006)
9. Sunstein, C.R.: Infotopia: How many minds produce knowledge. Oxford University Press, Oxford (2006)
10. Watts, D.J.: A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences of the United States of America 99(9), 5766–5771 (2002)