Angels or demons? Classifying desirable heavy users and undesirable power sellers in online C2C marketplace

Hikaru Yamamoto¹ · Nina Sugiyama² · Fujio Toriumi² · Hikaru Kashida³ · Takuma Yamaguchi³

Received: 31 July 2018 / Accepted: 17 July 2019 / Published online: 1 August 2019
© The Author(s) 2019

Abstract

To grow and succeed, online consumer-to-consumer (C2C) marketplaces need to increase the number of users and transactions, because their main revenue is usually the transaction fee. To increase the number of users and transactions, uncertainty must be reduced and a safe and enjoyable transaction environment must be maintained. In this paper, we aim to detect malicious users and power sellers who can harm the healthy growth of an online C2C platform. Using the data set of a major online C2C marketplace called Mercari, we classified undesirable users by building a classification model for banned users. The results of the banned user prediction indicated that most banned users are heavy sellers. Heavy sellers are desirable from the viewpoint of increasing the transaction fee revenue, but many are power sellers who are running full-time businesses on the platform, making it difficult for non-professional sellers to compete, and their dominance may eventually alienate users. Thus, we built another classification model to classify desirable and undesirable power sellers. Applying the model to the CART classifier, we successfully classified non-professional heavy users and undesirable power sellers in an online C2C marketplace.

Keywords Online platform · Consumer-to-consumer marketplace · Online fraud · Banned user prediction

Introduction

Online marketplaces are broadly categorized into two categories: (1) business-to-business (B2B) and (2) business-to-consumer (B2C) or consumer-to-consumer (C2C) [19]. A B2B marketplace facilitates exchange among organizations, whereas

✉ Hikaru Yamamoto
hikaru@kbs.keio.ac.jp

Extended author information available on the last page of the article

a B2C or C2C marketplace facilitates transactions involving consumers. This paper focuses on a C2C marketplace.

C2C transactions differ from B2C transactions in various ways and require new models of operation [13, 15, 18]. First, the degree of seller uncertainty is considerably larger in C2C transactions. According to Dimoka et al. [8], seller uncertainty is defined as the buyer's difficulty in assessing the seller's true characteristics and predicting whether the seller will act opportunistically. The quantity and quality of the information regarding sellers are limited on a C2C platform, whereas on a B2C platform, the sellers provide full information of their companies, products, and services. In some cases, platforms allow users to use a nickname and stay anonymous during the transaction process. The lack of information on users raises the uncertainty and risk of the transaction for both sellers and buyers. The second difference is the continuity of the buyer–seller relationship. The sellers in a C2C marketplace are not professional sellers, but consumers who have become amateur retailers [13]. This means that sellers are “open” for business only when they want to be. Furthermore, these amateur sellers rarely have an inventory of their products. Given that the products sold in C2C transactions are usually secondhand, the buyer–seller relationship tends to end in one transaction. In summary, C2C transactions are much more uncertain and riskier than B2C transactions; thus, malicious users must be detected for the platform to grow and succeed. In this paper, we aim to detect malicious users and power sellers who can harm the healthy growth of an online platform. Using the data set of a major online C2C platform, we classified undesirable users by building a classification model for banned users and applying it to a Classification And Regression Tree (CART) classifier. The results indicated that most banned users are heavy sellers. The next challenge is to classify desirable heavy sellers who are non-professional sellers that are purely heavy users of the service and undesirable power sellers who are running full-time businesses on the platform. Using the CART classifier, we successfully classified non-professional heavy users and undesirable power sellers in an online C2C marketplace.

Related work

Online fraud detection

Although most users tend to be honest, some may engage in online fraud activities. To maintain a safe transaction environment for the users, malicious users need to be detected and banned as early as possible. Many researchers have proposed ways to detect malicious users in online communities. For example, Akoglu et al. [2] proposed an unsupervised, general, network-based framework for opinion fraud detection. They defined the opinion fraud detection problem as a classification task on signed bipartite review networks. Cheng et al. [4] characterized antisocial behavior such as trolling in online discussion communities by analyzing users who were banned from the communities. Similarly, Li et al. [16] detected fake reviews in a large-scale data set using temporal and spatial features. Luca and Zarvas [17] empirically analyzed review fraud on Yelp and investigated the conditions under which it

is most likely to occur. Heydari et al. [14] reviewed literature on spam review detection published between 2007 and 2014 and observed that the most efficient technique is to capture burst patterns as spam attacks and investigate reviews that fit the patterns. They also categorized detection techniques as (1) finding duplicates and (2) content-based methods such as genre identification, detection of psycholinguistic deception, and text categorization.

Regarding C2C marketplaces, there is a rich body of related research into internet auctions. Chua and Wareham and Gregg and Scott list and describe internet auction fraud types, which are summarized in Table 1 [6, 10].

Grazioli and Jarvenpaa [11] claim that a comprehensive strategy to reduce the occurrence of internet deception must include deterrence, prevention, and detection. We label such a strategy as a proactive strategy. Some aspects of a proactive strategy are to authenticate the identity of the seller, determine the seller's reputation, detect illegal/inappropriate listings, and detect out-of-the-platform transactions. Having systems in place for authenticating user identity and determining seller reputation will make it difficult for malicious sellers to settle transactions on the platform. Seller reputation systems help establish a seller's reputation, reduce fraud, and build trust in the marketplace [3, 7, 10, 19]. Buyers assess the sellers' remarks before the transaction, and the feedback mechanism is useful to create price premiums for trustworthy sellers as returns on their reputation [19].

Illegal/inappropriate listings are often detected using text mining and machine-learning techniques [5, 20]. An out-of-the-platform transaction involves finding a seller or buyer on the platform and settling the transaction outside the platform to avoid paying the transaction fee. Given that this conduct decreases the revenue of the platform, out-of-the-platform transactions need to be detected and preempted.

In addition to proactive strategies, there are reactive strategies. These include a mutual rating system for buyers and sellers, a customer support center, and a buyer guarantee system. Pavlou and Gefen [19] focused on institutional mechanisms such as feedback features, escrow services, and credit card guarantees that

Table 1 Types of internet auction fraud

Shilling	Seller bids on own auction to drive up an item's price
Bid shielding/multiple bidding	Auction buyer places multiple bids using different identities, withdraws high bids subsequently, and purchases the item at a low bid
Misrepresentation	Seller deceives buyer about the true value of an item
Fee stacking	Seller adds hidden charges to the item's price after the auction ends
Failure to ship	Seller never sends the goods
Failure to pay	Buyer never sends the money
Reproductions and counterfeits	Seller advertises counterfeit goods as the real thing
Triangulation/fencing	Seller uses a stolen credit card to buy from an online merchant and resells the item at an auction
Buy and switch	Buyer receives merchandise but switches original merchandise with inferior merchandise before returning it
Loss or damage claims	Buyer claims the item was damaged and disposed of and requests money back
Shell auction	Seller sets up an auction solely to obtain names and credit card details

are implemented or created by a third party to create conditions that will facilitate successful transactions. They evaluated the perceived effectiveness of such mechanisms. Giving a negative rating to a malicious buyer or seller is effective when problems are encountered on the platform. The customer support center is the key to solving any problems arising during the process of C2C transactions. Finally, the buyer guarantee system can solve problems such as a buyer not receiving a product or receiving a fraudulent product.

Chicken-and-egg problem and power seller

The main revenue of C2C platforms is the transaction fee. The platform charges a certain percentage of the closing price from both the seller and the buyer. In most cases, a user can become a member of a C2C marketplace and products can be listed free of charge. C2C platforms need to attract both buyers and sellers and increase the number of transactions to be successful. These platforms exhibit two types of network effects, a same-side effect and a cross-side effect [9]. Given that the cross-side effect works strongly in the C2C marketplace, increasing the number of sellers is valuable for buyers, and vice versa. Although the notion of network effects is widely acknowledged, they are difficult to realize and enjoy because of a chicken-and-egg problem, to attain a critical mass of buyers, you need a critical mass of sellers—but to attract sellers, you need a lot of buyers.

To overcome this hurdle, it is tempting for an operator to depend on power sellers. Power sellers are defined as users who have moved from selling as a hobby or source of supplemental income to running a full-time business in the marketplace [12].

The reason that power sellers are attractive for a platform is that they cost-effectively increase the number of listings. However, growth through power sellers can result in undesirable consequences for the platform. According to Hagiu and Rothman [12], eBay discovered that the dominance of power sellers possibly hurts the buyer experience. For example, power sellers tend to favor “bulk listings”. However, bulk listings reduce the diversity of products for sale, resulting in crowding out unique products and causing the quality of the average listing to drop. Over the years, power sellers can become very powerful and come dominate a platform’s supply side. In such cases, it becomes difficult for non-professional sellers to compete, which will eventually alienate users.

Overview of C2C marketplace

Topic of analysis

The data were provided by Mercari, Inc. Mercari is the largest smartphone-based C2C marketplaces in Japan, which was released in July 2013, and has over 1 billion listings.

Mercari's free mobile app allows sellers to offer a wide variety of items, such as apparel, accessories, and games (Fig. 1). This app is a two-sided platform consisting of buyers and sellers. Its easy-to-use and intuitive interface allows easy, frictionless transactions between buyers and sellers. Users can browse products on Mercari for free and can buy or sell products after completing the user registration (date of birth, nickname, email address, password, phone number, address, and payment method). The user can become a seller by simply uploading the photo of the product along with a few details such as a title and description of the item. According to Mercari, the listing process takes approximately 3 steps and is free of charge.

Furthermore, Mercari offers an easy shipping experience. The app has a built-in shipping feature, so that the seller can drop the packaged product at a convenience store or post office without printing the label. When the transaction is completed, the buyer pays the price of the product for Mercari. Mercari takes a 10% transaction fee from the seller.

For platform governance, the company employs various tactics. To protect the users, the company provides an escrow service that holds the payment until the buyer confirms receipt of the purchased product. It also has a buyer guarantee system that covers the cost of purchase if the buyer does not receive the product. To maintain a healthy and accountable platform, the company uses a three-point-scale mutual rating system (good/normal/bad) for both buyers and sellers. In addition, for healthy growth for the platform and to maintain honest and safe transactions for users, Mercari detects and bans power sellers, described above.

Fig. 1 Screenshots of Mercari search results and seller profile

Analysis

Data overview

We were provided official user log data from April 2016 to December 2017. The data included the profile and transaction data for 109,158 banned users and 93,856 non-banned users. The banned users are detected by the company's machine-learning technique and inspected by staff at the customer support division. Thus, the banned user data set is handpicked "teacher" data and constitutes "ground truth" in our analyses. We use a commonly used supervised learning technique for detecting banned users. For each user, we were provided with the user profile (text), their purchases (date, item, product details), listings (date, item, product details), ban record (date, cause), and whether the user was invited to Mercari via a referral program.

Analysis 1, banned user prediction

To predict the banned users, we used the features in Table 2. All the features are continuous, except for the invitation dummy. Most features are self-explanatory, so we will explain some but not all.

When a seller lists an item, the seller posts a photograph of the item and uploads it with the title and description of the item, such as size and place of purchase. For *# of exact matches of listed item titles*, we investigate the listed items that have identical titles to other items listed by the same user and count the number of the exact matches. For *# of exact matches of listed item details*, the text of listed item descriptions is calculated in the same manner.

Table 2 Summary statistics of features for prediction

Features	Mean (<i>N</i> =203,014)
Length of profile text	125.1
# of purchases	8.4
# of listings	206.8
# of purchases/day	0.2
# of listings/day	2.6
# of same profile text	9.6
Entropy of listing time	1.4
Variance of listed product category	0.6
Variance of transaction day of the week	1.50
Time until first listing (day)	146.60
# of exact matches of listed item titles	154.10
# of exact matches of listed item details	183.40
# of listings in the first 3 days of membership	4.80
Buy/sell ratio (buy/buy + sell)	0.30

Table 3 Performance for banned user prediction

	Prediction, banned	Prediction, non-banned	Recall
Banned	9925	924	0.91
Non-banned	520	8933	0.94
Precision	0.95	0.91	0.93

Fig. 2 Visualization of banned user classification

Time until first listing (days) is defined as the # of days a user took to list the first item after creating a user account. The buy/sell ratio is defined as $(\# \text{ of purchases})/(\# \text{ of purchases} + \# \text{ of listings})$. The closer the value is to 0 (1), the more the user transacts as a seller (buyer). The mean for each feature is presented in Table 2.

Results

To predict banned users, we adopted a holdout validation strategy. Using the above-mentioned data set, we divided the data into training data (banned users 98,309, non-banned users 84,403) and test data (banned users 10,849, non-banned users 9453). We built a classification model and applied it to a CART classifier. Table 3 presents the performances of predicting banned users, and Fig. 2 presents the visualization of the classifier.

Understanding banned users

To understand which users are banned from a C2C platform, we illustrate how banned users differ from non-banned users. To do this, we interpret the visualization of the CART classifier, as shown in Fig. 2. The orange terminal nodes are groups of banned users, and the blue ones are groups of non-banned users. The results show that five groups of banned users can be classified with accuracy of 0.7 or above. The statistical summary of each group is presented in Table 4. The classification rules

Table 4 Summary Statistics of banned users

Features	Average	G1 (N = 2000)	G2 (N = 3493)	G3 (N = 2343)	G4 (N = 77,855)	G5 (N = 1089)
Length of profile text	125.1	50	22.1	24.9	196.2	348.2
# of purchases	8.4	0	0	10.2	0.2	40.5
# of listings	206.8	6.6	2.5	0.0008	404.2	1648.3
# of purchases/day	0.2	0	0	6.9	0.002	0.2
# of listings/day	2.6	0.9	1.3	2.90E-06	4.7	7.4
# of same profile text	9.6	8.5	5.4	3.1	19.3	5.8
Entropy of listing time	1.4	0.3	0.5	0	2.2	2.6
Variance of listed product category	0.6	0.1	0.2	0	0.7	0.9
Variance of transaction day of the week	1.50	0.6	0.8	1.1	1.6	1.9
Time until first listing (day)	146.60	9.3	13.4	281.6	19.3	20.2
# of exact matches of listed item titles	154.10	0.2	0.2	0	325.4	1200.2
# of exact matches of listed item details	183.40	0	1.5	0	388.3	1500
# of listings in the first 3 days of membership	4.80	1.2	0.8	0	9.4	12.3
Days of membership	224.00	13.2	30	91.4	117.3	254.9
Buy/sell ratio (buy/buy + sell)	0.30	0	0	1	0.004	0.04

for each banned group are as follows. The most important features are # of exact matches of the listed item details, # of purchases/day, time until first listing (day), buy/sell ratio, # of listings, and # of purchases.

Group 1

of exact matches of listed item details=0
of purchases/day ≤ 0.001
of listings/day > 0.04
time until first listing (day) > 0.541
variance of transaction day of the week ≤ 1.103

Group 2

of exact matches of listed item details=1 or 2
of purchases/day ≤ 0.001
of listings ≤ 3

Group 3

of exact matches of listed item details ≤ 2
of purchases/day > 0.001
time until first listing (day) ≤ 468.382
buy/sell ratio > 0.999

Group 4

of exact matches of listed product details > 2
of purchases ≤ 3

Group 5

of exact matches of listed item details ≥ 348
of listings > 357
of purchases > 3

We further investigate each banned group to understand the characteristics of banned users. Users in Group 1 are relatively new. Their days of membership are relatively short (13.2) and *# of listings* is low (6.6). This suggests that they were banned in the early stage of their membership for illegal listings. This argument also applies to Group 2, whose average days of membership is 30 days and *# of listings* is 2.5 items.

The users in Group 3 are not active sellers, since their *# of listings* is almost zero. They were members for about 3 months. We investigated the reasons for the bans, which were coded by the customer support center. The reasons were malicious activities such as creating multiple accounts, using invalid credit cards, and laundering money.

Group 4 is the largest group ($N=77,855$) accounting for 70% of the banned users. They are power sellers with large *# of listings*, *# of exact matches of listed item details*, and *# of exact matches of listed item titles*. Although they regularly listed

items, they almost never purchased items in the marketplace ($\# \text{ of purchase} \leq 3$). A seller needs to upload a title and description of an item when listing. Thus, there is a motivation a frequent seller of the same item to copy and paste the product description. The motivation is even stronger for a power seller who engages in bulk listings. The results confirm power sellers regularly reuse the same item titles and details in this way.

Users in Group 5 are also considered to be power sellers, since they have the largest $\# \text{ of listing}$ (1648.3). In addition, they seem to copy and paste the descriptions of listed items, because their $\# \text{ of exact matches of listed item details}$ is eight times larger than the average of users in the data set.

To summarize, there are two types of banned users, malicious users who are banned in the early stage of membership for illegal activities (Groups 1, 2, 3), and power sellers who are not engaging in online fraud but running full-time businesses on the platform (Groups 4, 5). Our results show that nearly, 80% of banned users are considered to be power sellers. While it is possible to observe that the majority of banned users are heavy sellers from simple descriptive statistics of the raw data, it is valuable to understand the detailed characteristics of the undesirable heavy sellers such as threshold value of buy/sell ratio and $\#$ of listings in the first 3 days of membership, as these features enable the platform to detect the undesirable users in the early stage.

Analysis 2, classification of heavy sellers

As explained earlier, the main revenue of C2C marketplaces is transaction fees, and Mercari takes 10% of the transaction price from the seller. Given that the transaction fee from the seller is the sole source of profit, heavy sellers can be considered as desirable users. The results of our first analysis indicate that the banned users are mostly heavy sellers. This suggests that there are two types of heavy sellers, i.e., angels with high monetary value who do not sell full time and demons who are banned for selling full time.

In marketing literature, both these types of users are considered as loyal customers with high monetary value. A frequently used analysis is RFM (recency, frequency, monetary) analysis, which ranks the customer by three aspects, the recency, frequency, and total monetary value of their purchases. In the case of a C2C marketplace, a sale can function as a purchase, since it generates a transaction fee. RFM variables provide the ground for sophisticated customer analysis. For example, Schmittlein et al. (1987) proposed a Pareto/NBD model that accounts for the relationship between recency and frequency and derived the probability of an individual customer being active at a particular point in time [21]. Abe (2009) further extended this model using a hierarchical Bayesian framework [1]. In the marketing literature, a customer who has high value for any of three RFM components is always desirable. However, this is not the case in the C2C marketplace, because the power sellers can negatively affect other non-professional buyers and sellers on the platform

This raises the importance of classification of the heavy sellers, i.e., classifying non-professional heavy sellers and power sellers who have moved from selling as a hobby or source of supplemental income to running a full-time business in the marketplace. In this section, we extract heavy sellers and build another classification model for desirable and non-desirable heavy sellers and apply it to a CART classifier. First, we define heavy seller as users whose *# of listings* ≥ 206 , as this value was the ground mean. Of the 35,946 users that met this condition, 90% (32,398) were banned users.

Table 5 shows the summary statistics of the heavy sellers. The summary statistics shows that the heavy sellers behave very differently from the average users in the marketplace. For example, their *length of profile text* is much longer than those for other users. In addition, their *buy/sell ratio* is much lower, suggesting that they act mainly as sellers. The heavy sellers also list items more frequently and start listing as soon, as they obtain the membership.

Results

Using a data set of 32,398 banned users and 3548 non-banned users, we divided the data into training data (banned users 29,174, non-banned users 3177) and test data (banned users 3224, non-banned users 371) to classify two types of heavy sellers. We again used the CART classifier and used the features in Table 5. The performances for classifying desirable and non-desirable heavy sellers are presented in Table 6. Since the CART classifier using our classification model classifies heavy sellers accurately enough, we proceed with interpreting its results.

Table 5 Features and summary statistics for heavy seller classification

Features	All	Heavy sellers
Length of profile text	125.1	266.8
# of purchases	8.4	10.2
# of listings	206.8	996
# of purchases/day	0.2	0.03
# of listings/day	2.6	8.9
# of same profile text	9.6	10.2
Entropy of listing time	1.4	2.6
Variance of listing product category	0.6	0.9
Variance of transaction day of the week	1.5	1.8
First listing (days)	146.6	22.8
# of exact matches of listed item titles	154.1	806.8
# of exact matches of listed item details	183.4	913.9
# of listings in the first 3 days of membership	4.8	17.6
Buy/sell ratio (buy/buy + sell)	0.30	0.02

Table 6 Performance for banned user prediction of heavy seller

	Prediction, banned	Prediction, non-banned	Recall
Banned	3146	78	0.98
Non-banned	82	289	0.78
Precision	0.97	0.79	0.96

Understanding desirable heavy sellers

In this section, we shed light on non-banned users. As explained earlier, heavy sellers are likely to be banned. However, non-banned heavy sellers are rare exceptions who are genuinely desirable for the growth of the marketplace. Figure 3 presents the visualization of the CART classifier. The blue nodes represent groups of non-banned user, and the orange nodes represent groups of banned users.

In Fig. 3, the blue terminal nodes in the bottom far right (148 banned users, 324 non-banned users), second from the left on the bottom (230 banned users, 500 non-banned users), in the bottom far left (314 banned users, 1562 non-banned users) are groups of desirable heavy sellers. To further understand the desirable heavy sellers, we illustrate their characteristics by interpreting the results of the CART classifier.

The first split is *# of exact matches of listed item details*. If *# of exact matches of listed item details* ≤ 172.5 , the user is likely to be desirable. The second split is *# of listings/day*. These features are important to classify power sellers from desirable heavy sellers (Figs. 4, 5).

Fig. 3 Visualization of heavy seller classification

Fig. 4 # of exact matches of listed item details

Fig. 5 # of listings/day

Conclusion and future research

In this paper, we classified banned users and illustrated their characteristics in online customer-to-customer (C2C) marketplaces. We built a classification model and applied it to a CART classifier (Analysis 1) to predict banned users. Its results

indicate that the banned users are mainly sellers in the marketplace.

Because the sole revenue of the marketplace is the transaction fees, heavy sellers can be considered desirable users. Our research suggests that there are two types of heavy sellers, i.e., angels with high monetary value who do not sell professionally and demons who are power sellers who are banned for running full-time businesses on the platform. To classify desirable and undesirable heavy sellers, we once again built a classification model and applied it to a CART classifier to understand the difference between the two.

Although our classification models are simple and utilize a commonly used technique, they make two significant contributions to C2C marketplaces. First, installing and improving a malicious user detection system is beneficial for the operator, because a safe transaction environment must be maintained for the marketplace to grow. Second, because the business growth depends on the growth of revenue from its main source (the transaction fee), it is important to encourage the users to list items. However, there are desirable and undesirable heavy sellers. Our classification of heavy sellers will contribute to exclude undesirable power sellers and detect and encourage desirable heavy sellers in the marketplace.

The future research includes sophisticating the classification models and testing other features that are available, but not included in the models. In addition, to further understand the characteristics of banned users, it may be worthwhile to analyze the evolution and growth of the users in the marketplace.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Abe, M. (2009). Counting your customers one by one, a hierarchical bayes extension to the pareto/NBD model. *Marketing Science*, 28(3), 541–553.
2. Akoglu, L., Chandy, R., & Faloutsos, C. (2013). Opinion fraud detection in online reviews by network effects. In *International AAAI Conference on Web and Social Media*. Retrieved from <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/5981>
3. Ba, S., & Pavlou, P. A. (2002). Evidence of the effect of trust building technology in electronic markets, price premiums and buyer behavior. *MIS Quarterly*, 26(3), 243–268.
4. Cheng, J., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2015). Antisocial behavior in online discussion communities. In *International AAAI Conference on Web and Social Media*. Retrieved from <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10469/10489>
5. Chiu, C., Ku, Y., Lie, T., & Chen, Y. (2011). Internet auction fraud detection using social network analysis and classification tree approaches. *International Journal of Electronic Commerce*, 15(3), 123–147.
6. Chua, C. E. H., & Wareham, J. (2004). Fighting internet auction fraud, an assessment and proposal. *Computer*, 37(10), 31–37.
7. Dellarocas, C. (2003). The digitization of word-of-mouth, promise and challenges of online feedback mechanisms. *Management Science*, 49(10), 1407–1424.
8. Dimoka, A., Hong, Y., & Pavlou, P. A. (2012). On product uncertainty in online markets: theory and evidence. *MIS Quarterly*, 36(2), 395–426.
9. Eisenmann, T. R., Parker, G., & Van Alstyne, M. (2006). Strategies for two-sided markets. *Harvard Business Review*, 84(10), 1–11.

10. Gregg, D. G., & Scott, J. E. (2006). The role of reputation systems in reducing on-line auction fraud. *International Journal of Electronic Commerce*, 10(3), 95–120.
11. Grazioli, S., & Jarvenpaa, S. L. (2003). Deceived, under target online. *Communications of the ACM*, 46(12), 196–205.
12. Hagiu, A., & Rothman, S. (2016). Network effects aren't enough. *Harvard Business Review*, 94(4), 65–71.
13. Hennig-Thurau, T., Malthouse, E. C., Friege, C., Gensler, S., Lobschat, L., Rangaswamy, A., et al. (2010). The impact of new media on customer relationships. *Journal of Service Research*, 13(3), 311–330.
14. Heydari, A., Ali Tavakoli, M., Salim, N., & Heydari, Z. (2015). Detection of review spam, a survey. *Expert Systems with Applications*, 42(7), 3634–3642.
15. Jones, K., & Leonard, L. N. K. (2008). Trust in consumer-to-consumer electronic commerce. *Information and Management*, 45(2), 88–95.
16. Li, H., Chen, Z., Mukherjee, A., Liu, B., & Shao, J. (2015). Analyzing and detecting opinion spam on a large-scale dataset via temporal and spatial patterns. In *International AAAI Conference on Web and Social Media*. Retrieved from <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10534>
17. Luca, M., & Zervas, G. (2016). Fake it till you make it, reputation, competition, and yelp review fraud. *Management Science*, 62(12), 3412–3427.
18. Moser, C., Resnick, P., & Schoenebeck, S. (2017). Community commerce, facilitating trust in mom-to-mom sale groups on facebook. *CHI'17*. <https://doi.org/10.1145/3025453.3025550>.
19. Pavlou, P. A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Information Systems Research*, 15(1), 37–59.
20. Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: implications for trust building, price premiums, and seller differentiation. *Information Systems Research*, 17(4), 392–414.
21. Schmittlein, D. C., Morrison, D. G., & Colombo, R. (1987). Counting your customers, who are they and what will they do next? *Management Science*, 33(1), 1–24.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Hikaru Yamamoto¹ · Nina Sugiyama² · Fujio Toriumi² · Hikaru Kashida³ · Takuma Yamaguchi³

Nina Sugiyama
sugiyama@g.ecc.u-tokyo.ac.jp

Fujio Toriumi
tori@sys.t.u-tokyo.ac.jp

Hikaru Kashida
hikaru@mercari.com

Takuma Yamaguchi
kumon@mercari.com

¹ Graduate School of Business Administration, Keio University, Yokohama, Japan

² Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

³ Mercari, Inc, Tokyo, Japan