"I Think You'll Like It": Modelling the Online Purchase Behavior in Social E-commerce

FENGLI XU, ZHENYU HAN, JINGHUA PIAO, and YONG LI*, Tsinghua University, China

Understanding the roles of social factors in online purchase behavior has been a long standing research problem. The recently emerging social e-commerce platforms leverage the stimulated word-of-mouth effect to promote the sales of items, which offers a peek into the complex interplay between the social influence and online purchasing behavior. In this paper, we investigate this problem on a full-scale purchase behavior dataset from one of the leading social e-commerce platforms, Beidian. Specifically, we conduct a comparison study between the social e-commerce and conventional e-commerce that are both integrated in Beidian to examine how social factors affect user's purchase behavior. We reveal that social e-commerce leads to a 3.09∼10.37 times higher purchase conversion rate compared with the conventional settings, which indicates users make purchase with significantly fewer item explorations. Then, we propose and validate four primary mechanisms that contribute to the efficient purchase conversion: *better matching*, *social enrichment*, *social proof* and *price sensitivity*. Moreover, we identify several behavioral indicators that are able to measure the effect of these mechanisms, based on which we design an accurate predictive model (AUC=0.7738) for user's purchase decision. These results combine to shed light on how to understand and model the purchase behavior in social e-commerce.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**; • **Applied computing → Electronic commerce**; **Sociology**;

Additional Key Words and Phrases: Social E-commerce; Recommender System; User Behavior

ACM Reference Format:
Fengli Xu, Zhenyu Han, Jinghua Piao, and Yong Li. 2019. "I Think You'll Like It": Modelling the Online Purchase Behavior in Social E-commerce. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 65 (November 2019), 23 pages. https://doi.org/10.1145/3359167

1 INTRODUCTION

What is the role of social influence in online purchases? As the e-commerce sites rapidly gain popularity in the past decade, extensive research efforts from both academia and industries have been attracted to this long-standing problem [26, 29, 56]. As a result, there have been several attempts in integrating social features into the e-commerce platforms, such as Amazon customer reviews, Facebook-driven F-commerce and Groupon. Different from these previous attempts, a newly-emerging form of social e-commerce platforms distinguishes themselves by leveraging the stimulated word-of-mouth effect (e.g. Beidian, Pinduoduo, Yunji). That is they provide financial incentives to motivate their users to recommend the platform's items to their friends via social

*This is the corresponding author

Authors' address: Fengli Xu, xfl15@mails.tsinghu.edu.cn; Zhenyu Han; Jinghua Piao; Yong Li, liyong07@tsinghua.edu.cn, Beijing National Research Center for Information Science and Technology (BNRist), Department of Electronic Engineering, Tsinghua University, Beijing, China.

2573-0142/2019/11-ART65 $15.00
https://doi.org/10.1145/3359167

networks. These platforms present a unique opportunity to examine the complex interplay between the social influence and user's purchase decisions.

We carry out our study with a full-scale purchase behavior dataset from one of the leading social e-commerce platforms in China, Beidian. It includes both the social and conventional e-commerce scenarios. That is the users not only can browse, add to cart and purchase items in a conventional e-commerce interface, but also can purchase items through the web links from their friend's recommendations. Therefore, it enables us to conduct an in-depth comparison analysis between the purchase behaviors in these two scenarios, which is vital to dissect the roles of social influence in online purchase behaviors.

Our preliminary analysis reveals that the social e-commerce scenario leads to a significantly higher purchase conversion rate (3.09~10.37 times higher than conventional scenario), which indicates users can identify their desired items with notably less item explorations. This observation might bring important insights for improving user experience and increasing the platform's revenue. Motivated by a deeper understanding of such purchase behavioral differences, our key research questions are: what are the primary mechanisms behind the significantly higher purchase conversion rate in social e-commerce? And can the influenced purchase behavior in the social e-commerce be accurately predicted?

Since this new form of social e-commerce mainly relies on the simulated word-of-mouth effect, it has important links with the studies on word-of-mouth marketing strategies [6], where customers are incentivized to bring in new customers. On the other hand, it exploits the propagation of social influences via social network, which is closely related to the literature on online social influence diffusion [8]. Through analyzing the previous works in these related areas, we propose four potential mechanisms that govern the purchase behavior in social e-commerce:

- *Better matching mechanism:* Compared with the algorithmic recommender systems in conventional e-commerce platform, the motivated users in social e-commerce are able to recommend items that better match the friend's interest in their social network.
- *Social enrichment mechanism:* The customer's affection toward the recommended items is enriched by her social relationship with the sellers, i.e., the motivated users who recommend the items.
- *Social proof mechanism:* Observing the friends' purchase on certain items increase the customer's propensity to purchase that items.
- *Price sensitivity mechanism:* The effect of the social influence is correlated with the price of the items, where the customers respond differently to the social influences when the items are of different price.

We conduct an empirical comparison study to evaluate the validity of the proposed mechanisms, and gain deeper understandings on how they take effect in real-world scenario. First, we uncover empirical evidences that supports the *better matching mechanism*. In social e-commerce, customers generally explore less items in each purchase session, and converge more rapidly to the purchased items in terms of item similarity via each exploration. It indicates the sellers indeed are able to better match their friend's preference, which help her more efficiently navigate to the items she want. Second, we reveal that both the click frequency and the purchase conversion rate are significantly biased toward each customer's most familiar seller. In addition, there is evidence that customers' propensity to purchase exhibits casual relationship with the actions of others within the social communities. These results are in accordance with the *social enrichment mechanism* and *social proof mechanism*. They indicate that both the social influence from the sellers and the social communities affect users' online purchase decisions. Third, as the *price sensitivity mechanism* predicts, we demonstrate the effect of social influence indeed is correlated with the prices of items.

More specifically, customer's purchases in the social e-commerce are more concentrated in items with relative low prices. In addition, the effect of *social enrichment mechanism* is more prominent in low price ranges.

Alongside with the data analysis, we identify several behavioral indicators that measure the effect of the proposed mechanisms. For example, the frequency and duration of interactions between the sellers and customers can serve as proxies to the social tie strength between them, which measures the effect of *social enrichment mechanism*. Such behavioral indicators allow us to trace and potentially leverage the effect of the mechanisms in the various applications in e-commerce. To demonstrate their effectiveness, we incorporate them as social feature into the classic recommendation models to predict customer's purchase decision, which typically rely on the collaborative filtering feature and the profile of users and items. Specifically, we train a collaborative filtering based recommendation model, i.e., filed-aware factorization machine (FFM) [32], and a content based recommendation model, i.e., random forest (RF) [39]. The results from both models show the social feature alone can achieve comparable performance with the classic recommendation models. More importantly, integrating the social feature into the recommendation models leads to significant performance boost, where the AUC increases from 0.7404 to 0.7738 for FFM ($p < 0.01$) and from 0.6509 to 0.7473 for RF ($p < 0.01$).

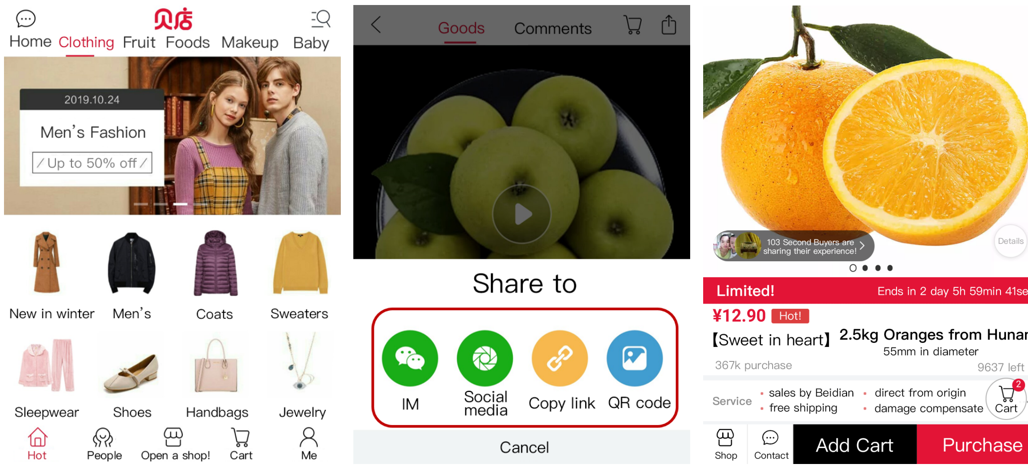
To conclude, the findings of this paper can be summarized into three points.

- We reveal the newly emerging form of social e-commerce leads to efficient purchase conversion, which is 3.09~10.37 times higher than conventional e-commerce scenario. In addition, we propose four primary mechanisms to rationalize these behavioral differences.
- We validate the proposed mechanisms with a large-scale comparison data analysis between the social and conventional e-commerce scenarios, which uncovers empirical evidences on how these mechanisms take effect.
- We identify several behavioral indicators to measure the effect of the mechanisms, on top of which we develop an accurate purchase model to predict users' purchase decisions in social e-commerce.

2 SOCIAL E-COMMERCE PLATFORM

To begin with, we take an overview of an emerging social e-commerce platform—Beidian. It is one of the largest platforms that enable the social e-commerce scenario by integrating the item recommendation feature into the application. As shown in Figure 1(a), this application provides conventional e-commerce service to users, which allows them to browse, add to cart and purchase various types of items. Moreover, it also facilitates users to recommend items to their friends in the social network via instant messages, social media and quick response codes (QR codes), which is shown in Figure 1(b). By clicking through the shared links, their friends will directly access the web pages of purchasing the recommended items (showing in Figure 1(c)). Therefore, each recommendation event constitutes a stimulated word-of-mouth recommendation that aims to leverage users' social connections to promote the items. Specifically, the platform provides commission fees for successful recommendations to motivate its users. To distinguish the roles of users, we use "sellers" to denote the users who initiate item recommendation, and "customers" to denote the recipients. In addition, we use "social e-commerce" to denote the scenario of purchasing the items thorugh their friend's recommendation, and "conventional e-commerce" to denote the scenario of purchasing items through the conventional e-commerce interface.

To examine the purchase behavioral differences between these two scenarios, we present their purchase conversion rates in Figure 2, by dividing the number of purchases with the number of clicks, which represents the efficiency of each item exploration of the users. From the results, we



(a) Interface of Beidian (b) Page of item recommendation (c) Page of recommended items

Fig. 1. Illustration of the Beidian platform's interface.

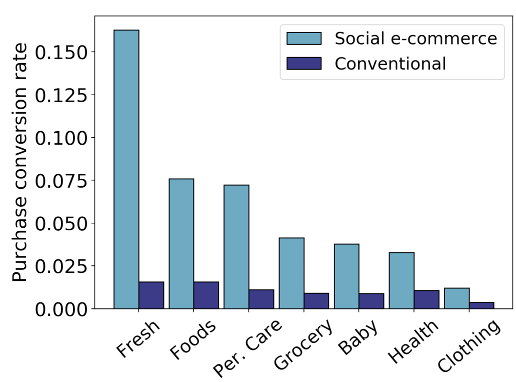


Fig. 2. The comparison between purchase conversion rate in social and conventional e-commerce scenarios.

can observe that the purchase conversion rate in social e-commerce scenario is notably higher than in conventional e-commerce scenario, and the differences vary across the item categories. Specifically, the smallest relative difference is in *Clothing* category (3.09 times higher), while the largest relative difference is in *Fresh Food* category (10.37 time higher). These observations suggest the purchase conversion is significantly more efficient in social e-commerce, which may facilitate the customers to identify their desired items more conveniently and generate more revenue for the platform. Therefore, it is of both academic and commercial interests to explore the underlying mechanisms behind these behavioral differences. These observations lead us to the first research question:

What are the primary mechanisms behind the significantly higher purchase conversion rate in social e-commerce?

Furthermore, to examine the effectiveness of the identified mechanisms in real-world applications, we aim to show case their implications by investigating the following question:

Can the purchase decisions in social e-commerce be accurately predicted?

3 BACKGROUND AND HYPOTHESIS

3.1 Social-aware Recommendation in E-commerce

Utilizing the social information to improve the recommendation performance has been a long-standing open problem that receives increasing attention [40]. The well known collaborative filtering paradigm leverages the idea of "familiar stranger" to recommend the purchased items of someone with similar purchase history to the users [51]. Furthermore, extensive research efforts are dedicated to improve the recommendation models based on the assumption of social homophily [45], where users who have social relations tend to exhibit more similar preference [50, 67]. On the other hand, another branch of researches focus on extracting more fine-grained recommendation signals from the social interface empowered by the platforms, such as the review function [65], instant message [23] and group event participation [66]. However, the previous research either focus on leveraging the social network information that is in parallel with the e-commerce platform or attempt to model spontaneous social interactions among the users. On the contrary, the core business model of the newly-emerging social e-commerce platforms is to incentivize their users to promote the items via their social connections. Therefore, it provides a novel and unique angle to examine the direct social influence the users will exert on their friends with explicit motivations, i.e. recommending or even persuading them to purchase certain items. Moreover, we aim to reveal the underlying mechanisms behind the efficient purchase conversion in social e-commerce, which haven't been adequately studied due to the lack of appropriate dataset.

Essentially, the social e-commerce platform requires the motivated sellers to act as "manual recommender system" to recommend items in a word-of-mouth manner, then the platform pays commission fees for their services in return. Such recommendations often take place on established social network, and may further propagate on them, e.g., retweeting. Therefore, this emerging form of e-commerce heavily relies on the stimulated word-of-mouth effect as well as the propagation of social influence. Thus, it has important links to the previous research efforts on word-of-mouth marketing strategies [6] and influence diffusion in social network [8]. We review the literature in these areas to develop key hypotheses for our research questions.

3.2 Word-of-Mouth Marketing Strategies

The power of word-of-mouth effect has been generally acknowledged by marketing practitioners and theorists [31]. Researchers find evidence that word-of-mouth is significantly more effective than advertisements from other sources in converting attitude [16], raising awareness [54] and influencing purchase decisions [42]. In addition, considerable research efforts have been dedicated to gaining more control over the word-of-mouth process, i.e., managing the stimulated word-of-mouth events. Following this idea, several active marketing strategies have been devised, which have two important branches: multi-level marketing [13] and referral marketing [4]. The multi-level marketing relies on cascading recruitment of sellers to form a "selling pyramid", and the profits are split level by level between the sellers and their "up line distributors" [61]. Despite of the resemblance in stimulating word-of-mouth recommendation with monetary reward, social e-commerce fundamentally differs from multi-level marketing strategy in the way that the profits are only split between the sellers and platform, i.e., no intermediary levels are involved. On the other hand, the referral marketing is based on incentivizing customers to bring in new customers with deliberately initiated and actively managed word-of-mouth referrals [59]. Therefore, the social e-commerce can be categorized as a type of online referral marketing program.

Previous research uncovers accumulating evidences that sophisticated referral marketing strategies often turn out to be profitable in numerous areas, including employee recruitment [14], service promotion [25] and product advertisement [24]. The customers acquired through referral marketing program tend to have longer retention period [62] and contribute higher profit margin [24], which effectively cover the marketing overhead. Several mechanisms have been developed to rationalize the efficiency of referral marketing, of which *better matching mechanism* and *social enrichment mechanism* have been considered as the most prominent ones [52]. Therefore, it is natural to assume these two mechanisms also play important roles in the high purchase conversion rate in social e-commerce. In the context of social e-commerce, *better matching mechanism* denotes the phenomenon that sellers' recommendations fit with the customer's preference better than the recommendations from other sources (i.e., algorithmic recommender systems in conventional e-commerce scenario). On the other hand, *social enrichment mechanism* indicates user's affection toward the items is enriched by their relationship with the sellers.

The monetary reward for successful recommendations is the most obvious driving force for sellers to make quality recommendations. In addition, the *better matching mechanism* also has root in several theories about the dynamics of social network. First, the reciprocity theory suggests sellers are likely to feel obliged to make suitable recommendations to their customers since they receive commission fees for the service [17]. There is evidence that referrers who receive commission fee generally acquired higher quality customers than others [18]. Second, the balance theory predicts user will recommend suitable items to their friends even without any incentives [30]. Finally, the social homophily effect indicates people tend to interact with others like themselves [45]. It implies the sellers are likely to have in-depth understanding on their customers, and hence are able to make recommendations in a tailored fashion to better meet their preference. Based on the above considerations, we make the hypothesis as follows,

H1 Better matching*: The purchase conversion rate in social e-commerce is higher because the sellers recommend items that better match the customer's preference.*

As for the *social enrichment mechanism*, it is built on top of the assumption that customer's perception of the items can be enhanced by their social bonds. For example, customer's trust toward a brand is strengthened if some closely related persons are also its customer [7]. In addition, customers acquired through their friends' referral generally show more commitment and attachment [52]. In the context of e-commerce, social enrichment mainly take effect in the form of trust [1, 29, 55]. Specifically, researches show that social relations are able to make customers feel more secure and increase the sales consequently [26, 57]. In addition, trust transfer theory have been proposed to depict the phenomenon that user's trust can be transferred from source to the target [44, 56]. This mechanism is also consistent with the prediction of the balance theory [30] and the social closure theory [60], where user's affection toward their friends is reflected to the item they endorse. To conclude, the *social enrichment mechanism* predicts the strength of social ties between the customers and the sellers contributes to the efficient purchase conversion. Therefore, we make the following hypothesis:

H2 Social enrichment*: Customer's propensity to purchase the recommended items positively correlates with her/his social tie strength with the sellers.*

3.3 Influence Diffusion in Social Network

With the prevalence of social media, the recommendations and purchases in social e-commerce are often made in public accessible platforms, e.g., micro-blogging services, group chatting services and online forums. Therefore, the influence of stimulated word-of-mouth often further propagates via social network, such as retweeting, which resembles the process of information cascading [9] and complex contagions [8]. Therefore, understanding the process of online influence diffusion may shed light on how social influence affect user's purchase behaviors. Previous works in this area mostly focus on predicting the structure of cascades in macro aspect [20] and the contagion behavior in a micro aspect [8]. The purchase conversion in social e-commerce is particularly related to the literature on complex contagions, since users are likely to be exposed to and affected by the actions of their friends. There is massive amount of evidence that individual's tendency to participate in a collective event often increases with the number of participants they have observed in their social communities, which is known as the "herd behavior" [12, 46]. Social proof theory is one of the prevalent narratives in rationalizing such behavior, which suggests users' decisions in ambiguous situations are often based on the believe that the majority's perspective is accurate [33]. Therefore, it is reasonable to assume the social proof plays an important role in customer's purchase decisions:

H3 Social proof: *A customer's propensity to purchase the items increases with the number of her/his friends who have already purchased them.*

On the other hand, the diffusion of purchase behavior fundamentally differs from copy-and-paste information diffusion cascades in terms of requiring considerable individual efforts. Specifically, in the context of e-commerce, the most apparent required effort is the cost of the items. Prior works on influence diffusion processes suggest the resource and effort needed for participating the online collective behavior play important roles in individual's decision making [43]. Users often only participate if the expected gains (e.g. improvement in social relationship and the utility of items) exceed the perceived costs (i.e. the price of items) [47]. As a result, the influence propagation with different individual effort requirement is observed to exhibit notably different patterns [11]. For example, the diffusion with less effort requirement (e.g. tweet resharing) often propagates more rapidly and forms large hubs, while the more effortful diffusion (e.g. ice bucket challenge) often propagates more slowly but more virally [11]. The well known elaboration likelihood model for persuasion provides a plausible reason for these previous findings [49]. That is the social influence affects user's behavior through the "central route" of rational appeal, where users often evaluate the potential social cost and required efforts as the pros and cons. Thus, in the context of social e-commerce, it is reasonable to assume that the effect of social influence is also affected by the required effort, i.e., the prices of the items. Therefore, we make the following hypothesis:

H4 Price sensitivity: *The effect of social influence, i.e., social enrichment and social proof mechanisms, is correlated with the price of items.*

4 COMPARISON ANALYSIS

In this section, we conduct a comparison analysis between the purchase behaviors in the social e-commerce and conventional e-commerce scenarios. We aim to evaluate the validity of the proposed hypotheses, and explore how they take effect in real-world scenario. We first introduce the datasets we collect, and then examine the proposed hypotheses point by point.

Table 1. The basic information of the collected dataset.

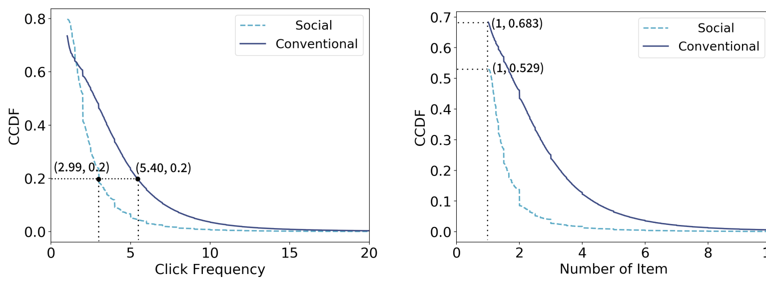
User	*Total*	*22,753,940*
	Active (Social)	16,351,927 (71.9%)
	Active (Conventional)	16,550,358 (72.7%)
	Active (Both)	10,155,838 (44.7%)
Item	*Total*	*419,342*
	Active (Social)	347,875 (83.0%)
	Active (Conventional)	412,304 (98.3%)
Category	First level	7
	Second level	256
	Third level	1,640

4.1 Dataset

We collect a full-scale dataset from both scenarios in Beidian platform for about four months, i.e., from 1 August to 27 November 2018. The dataset includes fine-grained click-through records, purchase records, user profile and item profile data. Specifically, the click-through records capture the events of customers clicking through the web links recommended by the sellers as well as the items they browse in conventional application interface. It consists of the customer's id, item's id, seller's id (only available in social scenario) and timestamp of the event. Similarly, the purchase records resemble click through record in data format, but capture the purchase events in both scenarios. In addition, the user profile data contains the gender, age and region information of user, while the item profile data contains the price and category of each item. The categories of items in Beidian are hierarchical and have three levels, with 7, 256 and 1,640 categories in each level accordingly. The basic statistic of dataset is presented in Table 1. Specifically, there are 22,753,940 users interacting with 419,342 items during this period. The records in conventional scenario cover 16,550,358 (72.7%) active users and 412,304 (98.3%) active items, while social scenario has 16,351,927 (71.9%) active users and 347,875 (83.0%) active items, where active denotes the users and items have at least one behavioral record.

It is worth to point out that the potential different user demographic distribution between the two e-commerce scenarios may pose challenge for rigorous comparison study. The reason is that the observational behavioral differences could be the result of the inherently different user population, which prevents us to attribute them to difference in e-commerce scenario. Fortunately, since both scenarios are integrated in one e-commerce platform, there are a large portion of users (10,155,838, 44.7%) that are active in both of them. This overlapped user population allow us to naturally control the user demographic distribution by limiting the comparison study on them. In this way, we can adequately eliminate any confounding factor that might be introduced by the different user population. Note that all the following comparison analyses between the two e-commerce scenarios are conducted on this overlapped user population, which allows us to draw solid conclusion on the behavioral differences between them.

Ethical considerations. To protect users' privacy, we take several procedures to eliminate the potential risk. First, the Terms of Service for Beidian include consent for research studies. Second, all potential individual identifiers are replaced with anonymous hashcode, so that no records can be linked to physical individuals. Third, all the research data is stored in an offline server, where the access is strictly limited to authorized researchers bound by confidentiality agreements.



(a) Number of clicks (b) Number of explored items

Fig. 3. The comparison of the required effort per purchase session between social and conventional scenario.

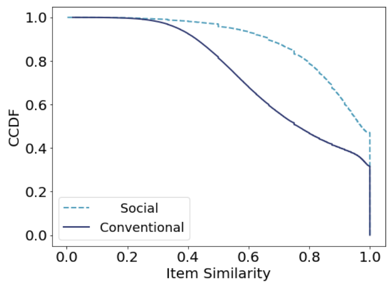
4.2 Better Matching Mechanism

Validating the *better matching* hypothesis requires to properly evaluate the fitness of each recommendation, which is difficult to measure with the purchase behavioral data. However, we can rely on the correlating behavioral indicators to shed light on user's satisfaction of recommendations. Specifically, the least effort theory [41] and information overload theory [22] suggest users tend to minimize their efforts in searching relevant information, which indicates less item explorations during the online purchase process generally leads to higher user satisfaction. Thus, we study number of clicks and explored items in each purchase process as the behavioral indicators of user satisfaction. Besides, to gain a more fine-grained understanding on how the item recommendation helps the users narrow down their search range, we also analyze the convergence of the item similarity during the item exploration of each purchase process.

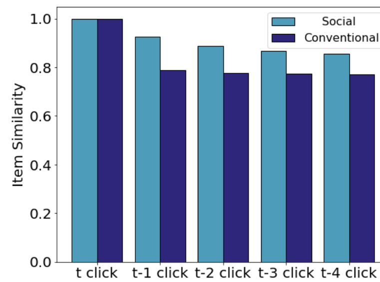
Social e-commerce reduces the effort of item exploration. We measure the quality of recommendation as the average effort users spend in each purchase process, which is similar with the effort-based user satisfaction metrics in recommender system evaluation [37]. The key idea is that if a user receive a quality recommendation then she might purchase the item without exploring others. Therefore, the less number of clicks and explored items in each purchase process indicates better recommendation quality. To quantify the number of clicks and explored items in each purchase process, we first segment the purchase records and click records into different purchase sessions with a purchase record or a 5 minutes inactivity period denote the end of each session.

Figure 3(a) and 3(b) depict the complementary cumulative distribution function (CCDF) of the number of clicks and explored item per purchase session in both e-commerce scenarios. From figure 3(a), we can observe that 80% of purchase sessions in social scenario require no more than 3 clicks, while the corresponding number in conventional scenario is 5.4 clicks. In addition, Figure 3(b) shows user's clicks are concentrated on a smaller range of items in social scenario. Specifically, 52.9% purchase sessions only explore 1 item in social scenario, compared with 68.3% purchase sessions in conventional scenario explore more than 1 item. In both figures, the curve of social scenario is consistently lower than conventional scenario, which indicates the general trend that users tend to click less and explore less items per purchase session in social scenario.

These results suggest users generally spend less effort in finding their desired items in social e-commerce, which supports the *better matching* hypothesis that the manual recommendations in social scenario outperform conventional algorithmic recommendations in matching user's preference. We also note that such phenomena might be the result of other correlating factors.



(a) Distribution of the item similarity between the purchased items and the explored items



(b) Variation of the item similarity between the purchased items and the explored items

Fig. 4. The distribution and variation of item similarity between the purchased items and explored items.

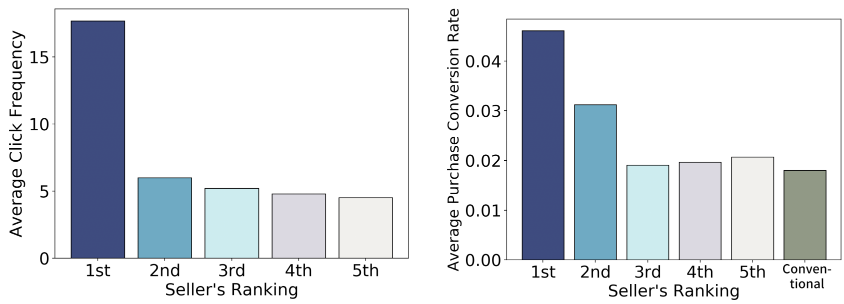
Mechanisms of social enrichment and social proof might persuade users to purchase items they actually do not want. In order to eliminate interference from other factors, we conduct a more fine-grained study on how user narrow down their search range with the help of recommendations.

Manual recommendations result in faster convergence to desired items.

The item profile data allows us to measure the similarity between the recommended items and the finally purchased items in one session, which serves as another indicator for the quality of recommendation. Besides, analyzing how the similarity varies across the explored items in each purchase session provides fine-grained understanding on the recommendation quality. It also allows us to distinguish the *better matching* mechanism from other social factors, since their effect should not vary much within short time period. Specifically, the similarity between two items is measured based on the category level they have in common, where the similarity score is evenly spaced between 0 and 1. For example, 1 means two items are identical, 0.75 means they share the same third level category (most fine-grained), and 0 means they are different in the first level category.

Figure 4(a) shows the CCDF of the similarity between the purchased items and the explored items within the same purchase sessions. We can observe the curve of social scenario is consistently higher than the conventional scenario, which means users generally explore the items that are more similar with the finally purchased items in social scenario. Moreover, Figure 4(b) shows the item similarity between the explored items and purchase items across different clicks, where t click indicates the final click that lead to purchase and $t-n$ click denotes the n-th click before it. We can observe the item similarity in social scenario is higher than that of conventional scenario across different clicks, and the social scenario has a higher starting point at the $t-4$ click. More importantly, there is a clear and steady convergence trend in social scenario, where the item similarity increases smoothly as the clicks approach the final purchase. It indicates the manual recommendation in social scenario can effectively help the users navigate to their desired items. On the contrary, the item similarity in conventional scenario does not vary much across different clicks. It suggests users in conventional scenario cannot effectively navigate based on the recommendations, but "stumble" into their desired items through inefficient item exploration. These results provide more fine-grained evidence that manual recommendation in social e-commerce effectively navigates users to their desired items iteratively, which conforms to the *better matching* hypothesis.

From the above analysis, we can conclude that both the click behavior pattern and the converging item similarity support the *better match* hypothesis. Seller's manual recommendations indeed



(a) Average number of clicks

(b) Average purchase conversion rate

Fig. 5. Customer's click frequency and purchase conversion rate with different sellers, where the sellers are ranked by their interaction frequency with the customers.

outperform the conventional algorithmic recommendation in terms of saving customer's effort for additional item exploration and helping them to more efficiently finding their desired items.

4.3 Social Enrichment Mechanism

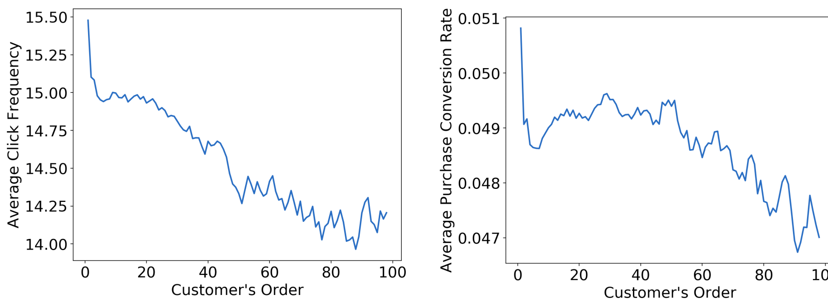
In the *social enrichment* hypothesis, we argue that the customers' propensity to purchase can be enriched by the social relation with sellers. Therefore, given a specific customer, the propensity to purchase might differ according to the sellers. That is the users might be more likely to purchase from the sellers they have close social relationship with, while less likely to purchase from those seldom contact ones. To evaluate the social tie strength between users, we use two widely adopted metrics [48], i.e., interaction frequency and duration of relationship. The idea is users tend to interact more with the people they are close to [21, 38], and the long duration of relationship is also a sign of strong social ties.

More interactions leads to more purchases.

To measure the social relationship strength with interaction frequency, we first compute the rankings of the sellers based on their interaction frequency with the given customers, where the top-1 ranking means that sellers have the stronger social relationships with the given customers. We show the average number of clicks and purchase conversion rates across the top-5 ranking sellers in Figure 5. Specifically, Figure 5(a) demonstrate a clear skewed click distributions, where the average number of clicks take place with the top-1 sellers is 2.95 times higher than that of the second sellers. In addition, the average number of clicks only has a slightly decrease from 5.99 to 4.52 from the second ranking sellers to the fifth ranking sellers. This observation suggests the customers are more likely to click the items recommended by their favorite sellers. As for the purchase conversion rate, Figure 5(b) shows the top-1 sellers have the highest purchase conversion rate of 0.046, which is 1.48 times higher than that of the second ranking sellers. Moreover, the purchase conversion rate gradually decreases to the conventional e-commerce level as the seller's ranking increases from 1 to 5. To conclude, our analysis reveals that both the click frequency and purchase conversion rate are significantly biased toward the customer's favorite sellers, which is consistent with the *social enrichment* hypothesis.

Longer the social relationship, higher the purchase conversion rate.

To measure the social tie strength as the relationship duration, we first order the customers of each user according to time their social relationship established. Then, the average number of



(a) Average number of clicks (b) Average purchase conversion rate

Fig. 6. The click frequency and purchase conversion rate the sellers receive from different customers, where the customers are ranked based on their relationship duration with the sellers.

clicks and purchase conversion rate are displayed in Figure 6. Specifically, Figure 6(a) shows the average number of clicks gradually drops from 15.48 to 14.10 as the ranking of customers increase from 1 to 80. It suggest the customers that have long social relationship with the sellers are more likely to click the recommended items. Similar observation is made in purchase conversion rate, which is displayed in Figure 6(b). The average purchase conversion rate also exhibit a decreasing pattern as the ranking of customer increases. Measuring the social tie strength with the relationship duration, we obtain another set of results that suggest the click frequency and purchase conversion rate increases with the social tie strength between the customers and sellers. These findings are consistent with previous analysis, and both support the *social enrichment* hypothesis.

The validity of the *social enrichment* hypothesis is evaluated with two widely adopt metrics from two different perspectives: from the customer's perspective, we measure the social tie strength based on their interaction frequency with the sellers; while from the seller's perspective, we measure the social tie strength as the relationship duration. Results from both perspectives show the propensity to purchase increase with the social tie strength between the sellers and customers. These two observations lead to the same conclusion that customer's propensity to purchase indeed is enriched by their social relationship with the sellers, which supports the proposed *social enrichment* hypothesis.

4.4 Social Proof Mechanism

The *social proof* mechanism describes the phenomenon that people in ambiguous situations tend to "copy" the decisions from their friends because they believe their friends possess more accurate information [33]. In this part, we evaluate the validity of the *social proof* mechanism in the online purchase behaviors in social e-commerce.

Similar purchase behavior within same community.

In the context of social e-commerce, *social proof* mechanism predicts the users within same communities should have more similar purchase behaviors than others due to the decision-copying phenomenon. To evaluate the accuracy of this prediction, we examine the purchase similarity among the user pairs within the same community and across different communities. Specifically, we measure the similarity as the JS-divergence [15] between the user's purchase frequency distribution on item categories, where smaller JS-divergence indicates more similar purchase behaviors. Figure 7 shows the comparison of the probability distribution function (PDF) of the purchase similarity

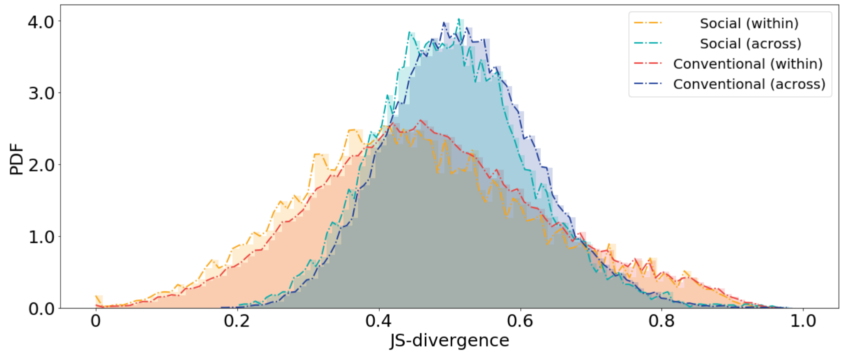
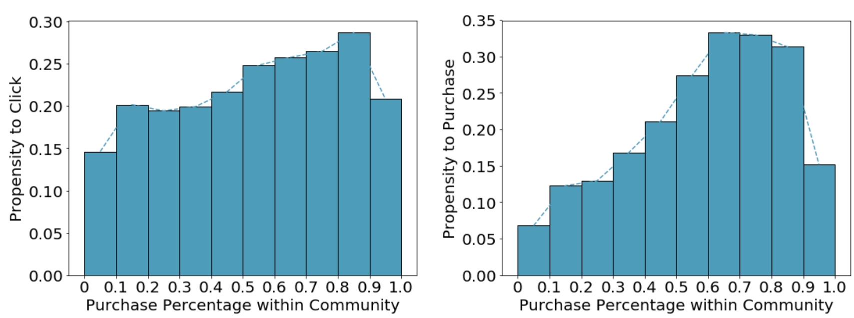


Fig. 7. The comparison of the purchase similarity among users within same community and across different communities in social and conventional scenarios.



(a) Propensity to click

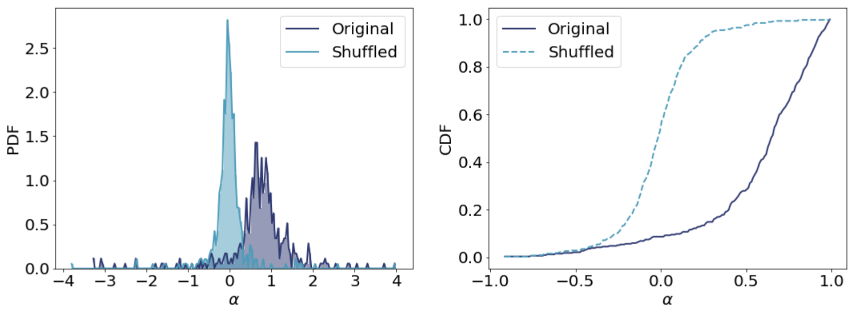
(b) Propensity to purchase

Fig. 8. The relations between the customer's reaction and their friend's purchase decisions.

within and across social communities, where the communities are defined as the customer groups that share at least one common seller. We observe that the user pairs from the same community generally have smaller JS-divergence compared with those from different communities. Moreover, the social scenario generally tends to have a smaller JS-divergence, which indicates the effect is more prominent in social e-commerce scenario. These observations are in accordance with the prediction of the *social proof* mechanism. However, we note that they can also be the results of social homophily effect [45], where humans tend to interact with the people that are similar to themselves and as a result the users within the same communities exhibit more similar purchase behaviors. To dissect the *social proof* mechanism from social homophily effect, we further carry out the following experiments.

Is it conformity or homophily?

To eliminate the interference of the social homophily effect, we aim to derive the causal relation between the customer's and their friends' purchase decisions. That is how the purchase decisions of one's friends directly impact on one's purchase decisions. Such analysis allows us to rule out the intrinsic similarity between the users. We capture the purchase decisions of one's friends as the percentage of customers within one's community that have already purchased that items. In addition, we measure the reaction of customers as the propensity to click and propensity to



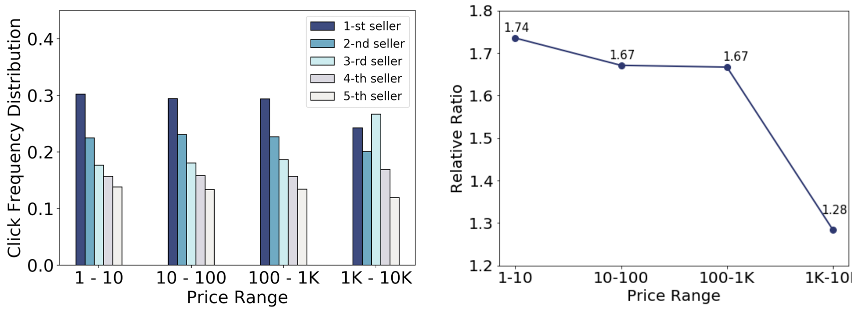
(a) Probability density function (b) Cumulative distribution function

Fig. 9. The distribution of α in original purchase behavior sequence and randomly shuffled sequence.

purchase, where the propensity to click is the percentage of customers within each community that click the items during one day and the propensity to purchase is the marginal purchase rate per day. We demonstrate the relations between the customer's reaction and their friend's purchase decisions in Figure 8. Combining both figures, we find the customer's propensity to click and purchase increases with the purchase percentage within their community before it reaches 0.8. These results indicate that the more customers within the same communities purchase the items, the more likely the other customers will be to click and purchase that items. The steady trend only has one exception when the purchase percentage exceeds 0.9, which is probably because there are a few "stubborn" users that are less likely to be affected by the *social proof* mechanism. These observations demonstrate the "herd behavior" in the social e-commerce, which is consistent with the prediction of *social proof* mechanism.

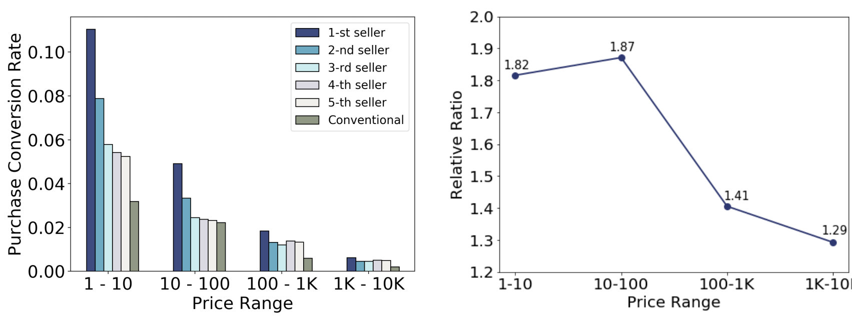
Furthermore, to more rigorously analyze the role of *social proof* mechanism in social e-commerce purchase behaviors, we leverage a well-established examine technique—*shuffle test* [2, 58]. A. Anagnostopoulos et al. [2] designed the shuffle test technique to appraise the casual relation in social influence with a variable α, where a large value of α indicates a high degree of causal relation. The key idea of shuffle test is that although user's behavior might be correlated because social homophily effect, the timing of their behaviors should be independent if there is no casual relation. Therefore, if the α in the original behavior sequence is significantly greater than the α in the randomly shuffled sequence, then there is evidence for the causal relation between users' behaviors. In the context of social e-commerce, we first order the purchase behavior records of a given item within each community into a time sequence based on the event time, and then we calculate the α of this sequence and the randomly-shuffled sequence. We demonstrate the distribution of the α variable in Figure 9. We can observe the average α of the original sequence is around 0.75, which is significantly greater than the average α in the shuffled sequence (two-tailed Student's t-test, $t = 16.5, p < 10^{-3}$). This result indicates there is a causal relation between the customers' purchase decisions within the same communities, which dissects the *social proof* mechanism from the social homophily effect and provides evidence for its existence.

Based on the above analyses, we discover the online purchase behaviors in social e-commerce are significantly influenced by the *social proof* mechanism, where there is a prominent casual relation between the customer's purchase decisions within the same community. In addition, our analysis shows the *social proof* mechanism positively contributes to the high purchase conversion rate,



(a) Distribution of the click frequency (b) Relative ratio of the click frequency

Fig. 10. The correlation between the click frequency distribution and the item price range.



(a) Distribution of purchase conversion rate (b) Relative ratio of purchase conversion rate

Fig. 11. The correlation between the purchase conversion rate and the item price range.

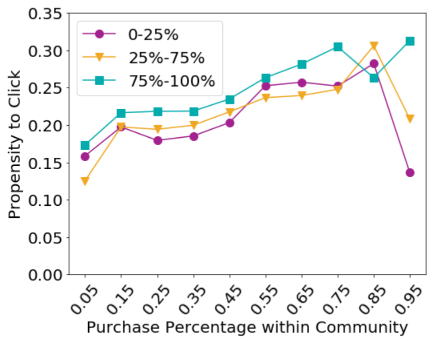
since customers tend to follow the purchase decisions of their friends. These results combine to support the *social proof* hypothesis.

4.5 Price Sensitivity Mechanism

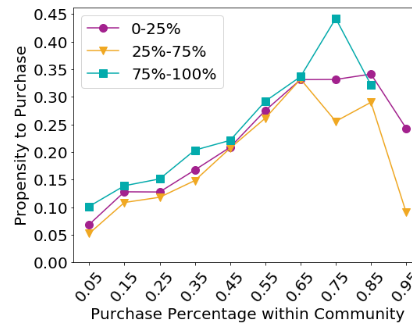
The *price sensitivity* hypothesis suggests customers might react differently to the social influence when the prices of the recommended items are different. Researchers have made similar findings which shows customers have different degrees of price sensitivity, which is mainly because different price perceptions [19], different degrees of brand loyalty [63, 64] and tariff awareness [36]. In this part, we intend to find out whether the effect of social influence, i.e., *social enrichment* and *social proof* mechanisms, is correlated with the price of the recommended items.

***Social enrichment* is more prominent in low price items.**

We examine the relation between the price range and effect of *social enrichment* mechanism. Specifically, Figure 10 and Figure 11 show the click frequency distribution and purchase conversion rate on different sellers across different price ranges. We can observe the click frequency and purchase conversion rate are both generally higher in low price range. More importantly, the bias toward the top-1 sellers is also more prominent in the low price range, and gradually decreases as the price range goes higher. To be specific, the relative ratio between the click frequency on the top-1 sellers and other sellers decreases from 1.74 to 1.28 as the price range increases from (1-10) to



(a) Propensity to click



(b) Propensity to purchase

Fig. 12. The correlation between the effect of *social proof* mechanism and the item price range.

(1k-10k), while the relative ratio of purchase conversion rate decrease from 1.82-1.87 to 1.29-1.41 as the price range increases from (1-100) to (100-10k). These observations demonstrate the effect of *social enrichment* mechanism plays a more important role in the items of low price range. It is probably because the customers are more cautious about purchasing the items of high price range, and less rely on their affection toward the sellers.

***Social proof* barely varies with item's price.**
In order to explore the relation between the item price and the effect of *social proof*, we show the curves of propensity to click and propensity to purchase on the items of different price ranges in Figure 12, where the items are classified into three groups of the top 25%, the bottom 25% and other items, respectively. Figure 12(a) shows similar trends of propensity to click on item of different price ranges, which are in consistent with the previous findings in *social proof* mechanism. Similar observations are made in the propensity to purchase curves in Figure 12(b). These results combined to suggest the item price does not play a significant role in the *social proof* mechanism. On the contrary to the *social enrichment* mechanism, the influence through *social proof* mechanism appears to be independent with the item price.

To sum up, the *price sensitivity* hypothesis holds in the social e-commerce, where customers respond differently to social influence on the items with different prices. More importantly, we find out that the differences mostly come from the *social enrichment* mechanism rather than from *social proof* mechanism. It will be an intriguing research quest to explore the underlying rationale. One plausible reason is that *social enrichment* mechanism can be categorized as a rational appeal where customers are more cautious about the high price items, while *social proof* mechanism is more likely to be an emotional appeal where customers pay less attention to the item itself [49].

5 PREDICTIVE MODELS

In the previous analysis, we propose and validate four primary mechanisms that lead to the highly efficient purchase conversion rate in the social e-commerce. To understand their implication for real-world applications, we aim to evaluate their effectiveness in predicting customer's purchase decisions in social e-commerce. Specifically, the task is predict whether a customer will purchase a certain item that is recommended by a specific seller or not. To reduce the interference of data sparsity, we first filter out the customers with less than 5 purchase records along with the items that are only available for less than three days. As a result, we derive a dataset consists of 89,550 users

Table 2. The feature sets selected for the predictive models.

Feature Set	**Selected Features**
Collaborative Filtering	user & item ids (for FFM model) user & item embeddings from Funk SVD (for RF model)
Profile	age, gender, number of purchase, region, city level (user profiles) item categories in three levels (item profiles)
Social	seller's ranking (*social enrichment*) interact frequency with the seller (*social enrichment*) number of friends who click the item (*social proof*) number of friends who purchase the item (*social proof*) price of the item (*price sensitivity*)
Combined	All features

and 99,055 items. We define the positive records as the recommendations that lead to successful purchase, and the negative records as the clicks without purchases. Furthermore, we split the dataset into a training set with 653,112 records (around 80%), validation data with 94,380 (around 10%) and testing data with 95,053 (around 10%). Since the customers have a smaller probability of purchasing the items, the number of positive records are significantly smaller than the negative records. We balance the training set by randomly down sample the negative records to the same amount of the positive records, which allows the prediction model to be trained in an unbiased manner. Note that the validation set and testing set are not balanced in order to evaluate the model's performance in the real-world scenario.

On the other hand, the setting of the predictive task is similar with the extensively studied recommendation problems, where there are two important branches of models — collaborative filtering models [32] and content-based model [39]. We evaluate the effectiveness of the developed mechanisms with one classic model from each branch, i.e., field-aware factorization machine (FFM) for collaborative filtering model [32] and random forest (RF) for content-based model [5]. These two models are chosen because they are the widely adopted recommendation models that can easily integrate additional heterogeneous features. To capture the effect of the identified social mechanisms, we first engineer the features that measure the behavior indicators of them. Specifically, we adopt the interaction frequency and seller's ranking defined in Section 4.2 to quantify the effect of *social enrichment* mechanism. For *social proof* mechanism, we calculate the number of friends who click and who purchase the same item within the same communities to represent its effect. Then, we model *price sensitivity* mechanism as the cost of the item. Since the *better matching* mechanism describes the general advantage of manual recommendation in social e-commerce compared with conventional scenario, it is not suitable to serve as features for predicting the purchase decisions. In addition to the social features, we also include the collaborative filtering (CF) feature set and the profile feature set about user and item profiles to serve as the important baselines of classic recommendation model. As for the CF feature set, we use the user ids and item ids in the FFM model since it automatically compute the collaborative filtering embedding vectors. On the contrary, we leverage the classic Funk SVD model [35] to compute user and item embedding vectors for RF model. Besides, we construct a profile information feature set containing all the available user and item profile features that are commonly used in content-based recommender

Table 3. Predicting purchase decisions in social e-commerce with different feature sets. We report the average results out of 20 experiments, and (_) denotes the standard deviation. ** indicates $p < 0.01$ significance level over all baselines with two-tailed Student's t-test.

Feature Set	Model	Precision	Recall	F1-score	AUC
CF	FFM	0.5966 (±0.0011)	0.6433 (±0.0027)	0.6191 (±0.0011)	0.7176 (±0.0010)
Profile	FFM	0.5615 (±0.0014)	0.6843 (±0.0046)	0.6168 (±0.0021)	0.6892 (±0.0020)
CF + Profile	FFM	0.6026 (±0.0035)	0.7110 (±0.0118)	0.6523 (±0.0055)	0.7404 (±0.0045)
Social	FFM	0.6094** (±0.0024)	0.7184 (±0.0054)	0.6594** (±0.0023)	0.7494** (±0.0024)
Combined	FFM	**0.6343**** (±0.0061)	**0.7293**** (±0.0131)	**0.6784**** (±0.0073)	**0.7738**** (±0.0068)
CF	RF	0.5367 (±0.0010)	0.5668 (±0.0017)	0.5514 (±0.0011)	0.6090 (±0.0004)
Profile	RF	0.5198 (±0.0004)	0.6151 (±0.0007)	0.5635 (±0.0005)	0.6171 (±0.0001)
CF + Profile	RF	0.5549 (±0.0006)	0.6501 (±0.0009)	0.5988 (±0.0007)	0.6509 (±0.0005)
Social	RF	0.5494 (±0.0004)	0.6271 (±0.0006)	0.5857 (±0.0004)	0.6741** (±0.0002)
Combined	RF	**0.6164**** (±0.0005)	**0.7344**** (±0.0008)	**0.6702**** (±0.0005)	**0.7473**** (±0.0003)

systems [39, 53], including item categories, age of user, gender of user and so on. Table 2 summarizes the feature sets we adopt in the predictive task.

For each evaluated model, we first tune the hyper parameter to the optimal with the validation set, then train 20 models on the training set and report their average performance on the test set along with the standard deviations. The classification threshold is set at 0.5 likelihood (ranges from 0 to 1), since the predictive models are trained on balanced training set [10]. The performance is measured with the metrics of precision, recall, F1-score and AUC, where the higher value indicates better performance for all the metrics. The results are showing in Table 3. We can observe that *CF* feature set and *Profile* feature set alone result in the lowest performances in both FFM model and RF model. On the other hand, the *Social* feature set identified by our proposed mechanisms outperform the *Profile* and *CF* feature sets in all performance metrics. Moreover, the *Social* feature set yields similar performance with the combination of *CF* and *Profile* feature sets, i.e., it performs slightly better in FFM model across all performance metrics and slightly worse in the precision, recall and F1-score in RF model. It indicates the *Social* feature set alone provides strong signals for predicting purchase behavior in social e-commerce, which is comparable with the standard recommender systems that leverage *CF + Profile* information.

Furthermore, we observe that the combination of all three feature sets generate the best prediction result in all performance metrics for both models. Specifically, it leads to significant performance gain over the standard recommender systems with *CF + Profile* features, which is 0.0714 in F1-score and 0.0964 in AUC for RF model ($p < 0.01$, two-tailed Student's t-test), while 0.0261 in F1-score and 0.0334 in AUC for FFM model ($p < 0.01$, two-tailed Student's t-test). We find that the FFM model generally outperform the RF model in *Combined* feature set, but the relative performance gain from incorporating *Social* feature is less prominent. One plausible explanation is that further improving the performance in FFM is more difficult. In addition, we also notice that the RF models generally perform worse than corresponding FFM models by 0.05, which is probably because most features are categorical and do not work well with RF model. However, this performance gap is narrowed down to less than 0.02 when the *Social* feature set is combined with the *CF + Profile* feature sets. It indicate the *Social* feature set serves as a good complement to the standard recommendation models, and integrating the *Social* feature set may result in a more robust performance. In conclusion, the experiment results combined to suggest the identified social mechanisms can benefit the real-world recommender systems, since they stand alone possess strong predictive power, and also work well to improve performance of standard recommendation models.

6 DISCUSSION

We conducted an empirical study on the purchase behavior of a social e-commerce platform. Specifically, we seek to reveal the underlying mechanisms behind the 3.09∼10.37 higher purchase conversion rate in this novel scenario. Through exploring the links with established social theories, we identify four potential mechanisms: *better matching*, *social enrichment*, *social proof* and *price sensitivity*. Furthermore, through extensive comparison analysis on the user behaviors between the social and conventional e-commerce scenarios, we evaluate the validity of the proposed mechanisms, and explore how they take effect in real-world applications. These findings have broad implications for both research communities and industrial practitioners, which can be mainly summarized as follow.

6.1 Social Influence Mechanisms in Online Purchasing

Understanding how social influences affect user's online purchase behavior has been a long standing research problem that attracts wide range of attentions [27, 28, 57]. A branch of closely related works focus on modelling the role of the social influence in establishing the trust for online purchasing [27, 34, 56]. For example, social influence models have been put forward to rationalize the trust-based item adoption [27] and word-of-mouth intentions [34]. However, most of these works study the social influence that is in parallel with the e-commerce platform and in a person-to-platform manner. We extend these previous findings on the novel scenario of social e-commerce, where users actively leverage their social influence to "sell" items to their friends in a person-to-person manner. We find out that the personal influence in this scenario can be characterized with *social enrichment* mechanism, where both the click through rate and purchase conversion rate significantly biased toward the sellers that have stronger social ties with customers. In addition, we complement previous studies with the findings on how community influences affect purchase behaviors. Specifically, we demonstrate *social proof* is another important mechanism in online purchasing. That is users tend to converge their actions to what the majority of their communities deem appropriate, in which conformity effect merges.

6.2 Improving Social-aware Recommendation

Previous works on social-aware recommendation mostly focus on how to leverage the available social information to improve the recommendation performance [65]. These models mainly build on top of the assumption that friends tend to share similar preference, which is known as social homophily [45]. For example, researchers demonstrate user's ratings on product review site tend to become more similar after they become friends [3]. As a result, numerous works are dedicated to harness its power by implementing "social reglarization" [40], which show promising results in performance improvement [50] and addressing cold start problem [67]. On the other hand, the full-scale behavior data collected from the social e-commerce platform allows us examine the fine-grained social influence in online purchase behavior. Therefore, we are able to identify four social mechanisms that go beyond the social homophily assumption. Our experiments show the behavior indicators of these identified social mechanisms exhibit strong predictive power, which can lead to significant performance gain in the standard recommendation models. Thus, our findings might shed light on the future design of social-aware recommender systems.

6.3 Integrating Social Features into E-commerce Platforms

Our empirical study finds evidences that properly integrating social features into e-commerce platform can lead to notable advantages. For example, the purchase conversion rate in social scenario is consistently higher than conventional scenario across all item categories. There are four primary mechanisms account for such behavioral differences: *better matching*, *social enrichment*, *social proof* and *price sensitivity*. These findings have direct implications for the attempts of integrating social features into e-commerce platforms.

The *better matching* mechanism suggests humans can do a better job than algorithms in recommending items to their social connections, which is probably because the fundamental limitations of algorithmic recommender systems, such as data sparsity and limited cognitive power. It indicates the performance of recommender system might be improved by keeping humans in the loop. For example, e-commerce platforms may implement a gamification system to motivate their users to identify the interest of their friends.

In addition, the *social enrichment* and *social proof* mechanisms imply user's purchase decisions are affected by social influence from their social connections and communities. To be specific, customers' propensity to purchase positively correlated with their social tie strength with the sellers. Moreover, users tend to mimic their friend's purchase decision. Therefore, both these two mechanisms contribute to the high purchase conversion rate. These results indicate the e-commerce platforms may promote the sales by reinforcing the effects of the social enrichment and social proof mechanisms. For example, they can provide convenient platform for users to exchange the purchase experiences with their friends. These findings may have important implications for the applications of promoting the e-commerce platform with social features.

6.4 Managing Word-of-Mouth Marketing

Our works also contribute to the large research bodies on word-of-mouth marketing strategies [4, 31]. Extensive research efforts have been devoted to actively manage the process of word-of-mouth marketing [59]. As a result, promising results have been reported in numerous areas, including customer acquisition [52] and service promotion [25]. Specifically, the mechanisms of *social enrichment* and *better matching* have been identified as the primary reasons of their success [52]. We study the problem of modelling the purchase behaviors in social e-commerce, which resembles the active word-of-mouth marketing in the way of motivating the users to perform stimulated word-of-mouth recommendations. In addition to the findings of previous research, we find out the *social proof* and *price sensitivity* mechanisms also play important roles in this novel scenario. Therefore, it is worthwhile to explore the implications of these findings in the context of word-of-mouth marketing.

6.5 Limitations

The research findings presented in our study are mainly derived from one social e-commerce platform, Beidian. Therefore, they might be biased due to the platform designs and the cultural background of the user population. However, the mechanisms identified in this study are based on the settings shared by similar platforms, which improves their generalization ability. In addition, because of the relative scarcity of such social e-commerce platform, it is hard to obtain secondary dataset to augment our findings. We are actively seeking the research opportunities on other platforms and leave the more comprehensive cross-platform analysis to future work.

Though our empirical comparison study does validate the overall effects of the proposed mechanisms, a more in-depth understanding on how these mechanisms play out in the wild require more fine-grained analysis on concrete purchase instances. Therefore, an important future work is to conduct a qualitative user study to further validate our findings, which may complement our research and allow us to gain deeper insights.

7 CONCLUSION

The recently emerging social e-commerce provides a unique opportunity to understand how social influences affect online purchase decisions. This work conducts a comparison study between the purchase behaviors in social and conventional e-commerce scenarios. We reveal that the purchase conversion rate in social scenario is 3.09~10.37 higher than conventional scenario. Furthermore, we propose and validate four primary mechanism behind the efficient purchase conversion: *better matching*, *social enrichment*, *social proof* and *price sensitivity*. We demonstrate that behavioral indicators for the effect of these mechanisms can be extracted from user behavioral data, which facilitates the design of an accurate purchase prediction model. These results provide a novel angle to integrate the social features into the various forms of e-commerce platforms.

ACKNOWLEDGMENTS

This work was supported in part by The National Key Research and Development Program of China under grant SQ2018YFB180012, the National Nature Science Foundation of China under 61971267, 61972223, 61861136003, and 61621091, Beijing Natural Science Foundation under L182038, Beijing National Research Center for Information Science and Technology under 20031887521, and research fund of Tsinghua University - Tencent Joint Laboratory for Internet Innovation Technology.

REFERENCES

[1] Hassan A Aljifri, Alexander Pons, and Daniel Collins. 2003. Global e-commerce: a framework for understanding and overcoming the trust barrier. *Information Management & Computer Security* 11, 3 (2003), 130–138.

[2] Aris Anagnostopoulos, Ravi Kumar, and Mohammad Mahdian. 2008. Influence and Correlation in Social Networks. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '08)*. ACM, New York, NY, USA, 7–15.

[3] Ching-man Au Yeung and Tomoharu Iwata. 2011. Strength of social influence in trust networks in product review sites. In *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 495–504.

[4] Barry Berman. 2016. Referral marketing: Harnessing the power of your customers. *Business Horizons* 59, 1 (2016), 19–28.

[5] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.

[6] Francis A Buttle. 1998. Word of mouth: understanding and managing referral marketing. *Journal of strategic marketing* 6, 3 (1998), 241–254.

[7] Emilio J Castilla. 2005. Social networks and employee performance in a call center. *American journal of sociology* 110, 5 (2005), 1243–1283.

[8] Damon Centola and Michael Macy. 2007. Complex contagions and the weakness of long ties. *American journal of Sociology* 113, 3 (2007), 702–734.

[9] Meeyoung Cha, Alan Mislove, and Krishna P Gummadi. 2009. A measurement-driven analysis of information propagation in the flickr social network. In *Proceedings of the 18th international conference on World wide web*. ACM, 721–730.

[10] Nitesh V. Chawla, Nathalie Japkowicz, and Aleksander Kotcz. 2004. Editorial: Special Issue on Learning from Imbalanced Data Sets. *SIGKDD Explor. Newsl.* 6, 1 (June 2004), 1–6.

[11] Justin Cheng, Jon Kleinberg, Jure Leskovec, David Liben-Nowell, Karthik Subbian, Lada Adamic, et al. 2018. Do Diffusion Protocols Govern Cascade Growth?. In *Twelfth International AAAI Conference on Web and Social Media*. 32–41.

[12] Robert B Cialdini and Robert B Cialdini. 2007. *Influence: The psychology of persuasion.* Collins New York. 13–31 pages.

[13] Brian Clegg. 2000. *The invisible customer: Strategies for sucessive customer service down the wire.* Kogan Page Publishers.

[14] James E Coverdill. 1998. Personal contacts and post-hire job outcomes: Theoretical and empirical notes on the significance of matching methods. *Research in Social Stratification and Mobility* 16 (1998), 247–270.

[15] Ido Dagan, Lillian Lee, and Fernando Pereira. 1997. Similarity-based Methods for Word Sense Disambiguation. In *Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics (EACL '97).* Association for Computational Linguistics, Stroudsburg, PA, USA, 56–63.

[16] George S Day. 1971. Attitude change, media and word of mouth. *Journal of advertising research* (1971), 31–40.

[17] Ernst Fehr and Simon Gächter. 2000. Fairness and retaliation: The economics of reciprocity. *Journal of economic perspectives* 14, 3 (2000), 159–181.

[18] Roberto M Fernandez, Emilio J Castilla, and Paul Moore. 2000. Social capital at work: Networks and employment at a phone center. *American journal of sociology* 105, 5 (2000), 1288–1356.

[19] Ronald E. Goldsmith and Stephen J. Newell. 1997. Innovativeness and price sensitivity: managerial, theoretical and methodological issues. *Journal of Product & Brand Management* 6, 3 (1997), 163–174.

[20] Benjamin Golub and Matthew O Jackson. 2010. Using selection bias to explain the observed structure of internet diffusions. *Proceedings of the National Academy of Sciences* 107, 24 (2010), 10833–10836.

[21] Mark Granovetter. 2018. *Getting a job: A study of contacts and careers.* University of Chicago press. 53–65 pages.

[22] Mary-Liz Grisé and R Brent Gallupe. 1999. Information overload: Addressing the productivity paradox in face-to-face electronic meetings. *Journal of Management Information Systems* 16, 3 (1999), 157–185.

[23] Stephen Guo, Mengqiu Wang, and Jure Leskovec. 2011. The role of social networks in online shopping: information passing, price of trust, and consumer choice. In *Proceedings of the 12th ACM conference on Electronic commerce.* ACM, 157–166.

[24] Sunil Gupta and Valarie Zeithaml. 2006. Customer metrics and their impact on financial performance. *Marketing science* 25, 6 (2006), 718–739.

[25] Alya Guseva. 2008. *Into the red: The birth of the credit card market in postcommunist Russia.* Stanford University Press. 1291–1293 pages.

[26] Mahmood Hajli. 2012. An integrated model for e-commerce adoption at the customer level with the impact of social commerce. *International Journal of Information Science and Management (IJISM)* (2012), 77–97.

[27] Mahmood Hajli. 2013. A research framework for social commerce adoption. *Information Management & Computer Security* 21, 3 (2013), 144–154.

[28] M Nick Hajli. 2014. A study of the impact of social media on consumers. *International Journal of Market Research* 56, 3 (2014), 387–404.

[29] BO Han and John Windsor. 2011. User's willingness to pay on social network sites. *Journal of computer information systems* 51, 4 (2011), 31–40.

[30] Fritz Heider. 2013. *The psychology of interpersonal relations.* Psychology Press. 59–79 pages.

[31] Paul M Herr, Frank R Kardes, and John Kim. 1991. Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective. *Journal of consumer research* 17, 4 (1991), 454–462.

[32] Yuchin Juan, Yong Zhuang, Wei-Sheng Chin, and Chih-Jen Lin. 2016. Field-aware factorization machines for CTR prediction. In *Proceedings of the 10th ACM Conference on Recommender Systems.* ACM, 43–50.

[33] Daniel Katz and Robert L Kahn. 1978. *The social psychology of organizations.* Vol. 2. Wiley New York. 257–333 pages.

[34] Sanghyun Kim and Hyunsun Park. 2013. Effects of various characteristics of social commerce (s-commerce) on consumersâĂŹ trust and trust performance. *International Journal of Information Management* 33, 2 (2013), 318–332.

[35] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 8 (2009), 30–37.

[36] Anja Lambrecht, Katja Seim, and Bernd Skiera. 2005. Does Uncertainty Matter ? Consumer Behavior Under Three-Part Tariffs. 1–4.

[37] Ting-Peng Liang, Hung-Jen Lai, and Yi-Cheng Ku. 2006. Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. *Journal of Management Information Systems* 23, 3 (2006), 45–70.

[38] Nan Lin, Paul W Dayton, and Peter Greenwald. 1978. Analyzing the instrumental use of relations in the context of social structure. *Sociological Methods & Research* 7, 2 (1978), 149–166.

[39] Pasquale Lops, Marco De Gemmis, and Giovanni Semeraro. 2011. Content-based recommender systems: State of the art and trends. In *Recommender systems handbook.* Springer, 73–105.

[40] Hao Ma, Haixuan Yang, Michael R Lyu, and Irwin King. 2008. Sorec: social recommendation using probabilistic matrix factorization. In *Proceedings of the 17th ACM conference on Information and knowledge management.* ACM, 931–940.

[41] Jens-Erik Mai. 2016. *Looking for information: A survey of research on information seeking, needs, and behavior.* Emerald Group Publishing. 55–79 pages.

[42] W Glynn Mangold. 1988. Use of commercial sources of information in the purchase of professional services: what the literature tells us. *Journal of Professional Services Marketing* 3, 1-2 (1988), 5–17.
[43] Gerald Marwell, Pamela E Oliver, and Ralph Prahl. 1988. Social networks and collective action: A theory of the critical mass. III. *Amer. J. Sociology* 94, 3 (1988), 502–534.
[44] D Harrison McKnight, Vivek Choudhury, and Charles Kacmar. 2002. Developing and validating trust measures for e-commerce: An integrative typology. *Information systems research* 13, 3 (2002), 334–359.
[45] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 1 (2001), 415–444.
[46] Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara, and Sune Lehmann. 2017. Evidence of complex contagion of information in social media: An experiment using Twitter bots. *PloS one* 12, 9 (2017), 1–14.
[47] Pamela Oliver. 1984. " If You Don't Do it, Nobody Else Will": Active and Token Contributors to Local Collective Action. *American sociological review* (1984), 601–610.
[48] Andrea Petróczi, Tamás Nepusz, and Fülöp Bazsó. 2007. Measuring tie-strength in virtual social networks. *Connections* 27, 2 (2007), 39–52.
[49] Richard E Petty and John T Cacioppo. 1986. The elaboration likelihood model of persuasion. In *Communication and persuasion*. Springer, 1–24.
[50] Dimitrios Rafailidis and Fabio Crestani. 2016. Joint collaborative ranking with social relationships in top-n recommendation. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, 1393–1402.
[51] Badrul Munir Sarwar, George Karypis, Joseph A Konstan, John Riedl, et al. 2001. Item-based collaborative filtering recommendation algorithms. *Www* 1 (2001), 285–295.
[52] Philipp Schmitt, Bernd Skiera, and Christophe Van den Bulte. 2011. Referral programs and customer value. *Journal of marketing* 75, 1 (2011), 46–59.
[53] Fabrizio Sebastiani. 2002. Machine learning in automated text categorization. *ACM computing surveys (CSUR)* 34, 1 (2002), 1–47.
[54] Jagdish N Sheth. 1971. Word-of-Mouth in Lov Risk lnnovations. *Journal of Advertising* (1971), 15–18.
[55] Dong-Hee Shin. 2010. The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with computers* 22, 5 (2010), 428–438.
[56] Katherine J Stewart. 2003. Trust transfer on the world wide web. *Organization Science* 14, 1 (2003), 5–17.
[57] Gayatri Swamynathan, Christo Wilson, Bryce Boe, Kevin Almeroth, and Ben Y Zhao. 2008. Do social networks improve e-commerce?: a study on social marketplaces. In *Proceedings of the first workshop on Online social networks*. ACM, 1–6.
[58] Jie Tang, Sen Wu, and Jimeng Sun. 2013. Confluence: Conformity Influence in Large Social Networks. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '13)*. ACM, New York, NY, USA, 347–355.
[59] Mirjam A Tuk, Peeter WJ Verlegh, Ale Smidts, and Daniel HJ Wigboldus. 2009. Sales and sincerity: The role of relational framing in word-of-mouth marketing. *Journal of Consumer Psychology* 19, 1 (2009), 38–47.
[60] Christophe Van den Bulte. 2011. Opportunities and challenges in studying customer networks. In *The connected customer*. Routledge, 25–54.
[61] Peter J Vander Nat and William W Keep. 2002. Marketing fraud: An approach for differentiating multilevel marketing from pyramid schemes. *Journal of Public Policy & Marketing* 21, 1 (2002), 139–151.
[62] Thorsten Wiesel, Bernd Skiera, and Julian Villanueva. 2008. Customer equity: an integral part of financial reporting. *Journal of Marketing* 72, 2 (2008), 1–14.
[63] Kristof De Wulf, Gaby OdekerkenâĂŘSchrÃűder, Frank Goedertier, and Gino Van Ossel. 2005. Consumer perceptions of store brands versus national brands. *Journal of Consumer Marketing* 22, 4 (2005), 223–232.
[64] Kanghyun Yoon and Thanh V. Tran. 2011. Capturing consumer heterogeneity in loyalty evolution patterns. *Management Research Review* 34, 6 (2011), 649–662.
[65] Ting Yuan, Jian Cheng, Xi Zhang, Qingshan Liu, and Hanqing Lu. 2015. How friends affect user behaviors? An exploration of social relation analysis for recommendation. *Knowledge-Based Systems* 88 (2015), 70–84.
[66] Shuo Zhang and Qin Lv. 2018. Hybrid EGU-based group event participation prediction in event-based social networks. *Knowledge-Based Systems* 143 (2018), 19–29.
[67] Tong Zhao, Julian McAuley, and Irwin King. 2014. Leveraging social connections to improve personalized ranking for collaborative filtering. In *Proceedings of the 23rd ACM international conference on conference on information and knowledge management*. ACM, 261–270.

Received April 2019; revised June 2019; accepted August 2019