ABSTRACT
In online marketplaces, an increasing number of producers depend on search and recommendation systems to connect them with consumers to make a living. In this talk, we discuss how these systems will need to evolve from the traditional formulations by incorporating the producer value into their objectives. Jointly optimizing the ranking functions behind these systems on both consumer and producer values is a new direction and raises many technical challenges. To overcome these, we lay out an end-to-end solution and present the results of applying this solution on Facebook Marketplace.

1 INTRODUCTION
With the rise of the creator and gig economies and the popularity of online marketplaces fostering them in recent years, media (from short-form entertaining videos to yoga lessons to political blogs), products (from food to apparels to electronics) and services (such as, ride hail, food delivery and rental) are created, sold and provided by a large number of creators, sellers and providers. At the heart of these marketplaces, search engines and recommendation systems are the primary distribution channels connecting consumers and producers. These systems not only help the consumers discover relevant and high-quality content, products or services but also play a critical role in the producer’s success (or failure) and long-term retention (i.e., continuing to produce on these marketplaces).

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2021 Copyright held by the owner/author(s).
https://doi.org/10.1145/3404835.3464924

2 A/B TESTING ON PRODUCER SIDE
To demonstrate the challenge of producer-side A/B testing, let us consider a simple test in which the items from the new producers in the test group are boosted. Because of the boost, these new producers get a better experience compared to the ones in the control group. However, when the boosting is ramped up to 100% (i.e., to all new producers), the items in the original test group will be ranked lower because significantly more items are getting boosted now. Thus, the impact when experimented at 1% is artificially inflated, and the producer-side A/B test result is incorrect.

The root cause of this is a violation of "stable unit treatment value assumption" (SUTVA) since the experience of the producers in either group also depends on the ranking model applied to the other producers outside the group. Given this, we develop a novel
producer-side A/B testing framework based on a counterfactual property: the items in the treatment are ranked at where they would be if the treatment is ramped to 100% of the producers. Similarly, the items in the control are ranked where they would be if the status quo is applied to all. Thus, the difference between treatment and control is independent of what applies to the rest, i.e., satisfies SUTVA. The readers can find details in our previous work [2].

3 OPTIMIZE FOR PRODUCER SUCCESS

Commercial ranking systems are usually trained from historical engagement data. As a result, they tend to concentrate on the items from a small set of producers and do not give a fair distribution of the traffic to the majority of other producers. Thus, these producers with less distribution usually do not have the success they expect. Alternatively, we can consider the incremental value of boosting on the item success defined as having a minimum number of cta or clicks. The incremental value formula: the items in the two groups, the test group is receiving an increasing number of transactions over time relative to the control group after initially having a similar number of transactions. The increase in transactions is interesting because by incorporating the seller-side objective, we actually deviate from the original rankings purely optimized for transactions. The reason behind this is two fold: (i) the increase in seller retention and listings (ii) the exploration effect of increasing the distribution for items and sellers that normally rarely get exposed.

4 FROM SUCCESS TO RETENTION

To empirically evaluate the optimization strategy, we run both buyer- and seller-side A/B tests. The buyer-side test is simply a standard A/B test which measures the short-term impact on buyer-side experience. The seller-side test is run on the counterfactual framework described in Section 2. Only the items of the sellers in the test group could get up-ranked, but they are positioned as if the reranking strategy in Section 3 was applied to all items. At the same time, the items of the sellers in the control group are ranked at the positions as if no reranking was applied at all. The empirical result from the seller-side A/B test shows a significant improvement of 10% on seller success (i.e., the number of unique sellers having at least one cta on a given day) while the buyer-side test reveals neutral results on short-term buyer experience.

Distributing success to more unique sellers on a marketplace would clearly improve sellers’ experience and be an important objective by itself. However, we also want to examine the impact of the seller experience on seller retention and the long-term growth of the marketplace. To measure these, we run a seller-side backtest over more than six months as of this writing. In the backtest, we observe that the seller success improvement translates to a significant increase in seller retention and the number of posted listings. Comparing the number of transactions received by the sellers in the two groups, the test group is receiving an increasing number of transactions over time relative to the control group after initially having a similar number of transactions. The increase in transactions is interesting because by incorporating the seller-side objective, we actually deviate from the original rankings purely optimized for transactions. The reason behind this is two fold: (i) the increase in seller retention and listings (ii) the exploration effect of increasing the distribution for items and sellers that normally rarely get exposed.

5 CONCLUSIONS

With the rise of the creator and gig economies, search and recommendation systems play an increasingly important role for producers and the society at large. In this proposal, we lay out a vision in which these systems will evolve by incorporating the impact on the producer side into the objective function. We also present our end-to-end solution to the technical challenges emerging with this new direction as well as the initial results on Facebook Marketplace.

REFERENCES