Analysis of Rewards on Reward-Based Crowdfunding Platforms

Yusan Lin∗ and Wang-Chien Lee‡
Department of Computer Science and Engineering
Penn State University
University Park, Pennsylvania 16802
yusan@psu.edu*, wlee@cse.psu.edu‡

Chung-Chou H. Chang
Department of Medicine
University of Pittsburgh
Pittsburgh, Pennsylvania 15213
changj@pitt.edu

Abstract—Today, crowdfunding has emerged as a popular means for fundraising. Among various crowdfunding platforms, reward-based ones are the most well received. However, to the best knowledge of the authors, little research has been performed on rewards. In this paper, we analyze a Kickstarter dataset, which consists of approximately 3K projects and 30K rewards. The analysis employs various statistical methods, including Pearson correlation tests, Kolmogorov-Smirnov test and Kaplan-Meier estimation, to study the relationships between various reward characteristics and project success. We find that projects with more rewards, with limited and late-added rewards are more likely to succeed. We also categorize and automatically annotate rewards into fifteen reward item types. We further analyze how different types of reward items are adopted across various categories of projects as well as how they potentially affect project success. We discover that projects with include the previous rewards are most likely to succeed. Also, different categories of projects may employ different best strategies of adopting reward item types to achieve pledging goals. Finally, we verify the efficacy of reward-related information through predicting project success. The result shows that features extracted from rewards help better predict the successes of crowdfunding projects.

I. INTRODUCTION

In recent years, crowdfunding has emerged as a new means of fundraising for social causes, entrepreneurial ventures, and for-profit businesses [1]. To meet the need, different types of crowdfunding platforms have surfaced, ranging from the traditional lending-based (e.g., Funding Circle) and equity-based (e.g., Invesdor), to the newer reward-based (e.g., Kickstarter) and donation-based (e.g., GoFundMe) crowdfunding platforms [2], [3]. Among them, reward-based platforms are the most popular [4]. On a reward-based platform, the creators offer a list of rewards at different prices, called menu pricing. The backers evaluate the rewards on the list (as incentives) to decide their support of the project. Once decided, the backers fund the project with amounts set on their selections of rewards. Finally, at a later time, the project creators deliver the promised rewards. The process is illustrated in Figure 1.

Rewards are the media for connecting backers and projects. We argue that rewards play an important role determining whether the projects eventually succeed. Time to time, the appeal of rewards may override the appeal of projects themselves due to the items (prizes) included in the rewards. For example, a user might not be interested in a project that makes creative items due to the appeal of rewards may override the appeal of projects themselves. However, if one of the rewards offers a 50% discount backing price for three chargers, the user may possibly back the project if she sees it as a good deal. Or, a reward that will list the backers’ names on social media to show appreciation may attract backers who find receiving publicity rewarding. Moreover, if a reward limits the offering of a special item, e.g., a charger of rose gold color, to only the first ten backers, it may raise some backer’s desire to support the project. In deed, based on our data analysis, we find that rewards with limited offerings usually achieve higher ratios between raised fundings and the targeted goals, whether these projects succeed or not, as shown in Figure 2.

Abundant studies have been conducted to understand how various aspects of projects, such as the project descriptions [5], project updates [6], timing of donations [7], and social mentions of projects [8], affect their successes, both quantitatively [3], [9] and qualitatively [10]. While there are hints about the important role rewards may play on reward-based crowdfunding platforms, to the best knowledge of the authors, there exists no data analysis on rewards in crowdfunding projects to understand their impacts on project success. Therefore, in this paper, we aim to analyze the rewards’ role on projects in reward-based crowdfunding platforms.

To proceed, we intend to analyze the explicit features of project rewards, such as number of rewards in a project, pricing, limited offers, and so on, to better understand their relationships to project success. Using real reward data collected from Kickstarter projects, we perform a series of statistical tests as appropriate on those explicit features, including Pearson correlation estimation between features, Kolmogorov-Smirnov test of equal distributions, Kaplan-Meier estimation on reward sell-out rates, and log-rank test of equal sell-out rates. In addition, by examining reward descriptions, we observe that items offered in rewards often exhibit similar characteristics, e.g., discount, appreciation, limited offering, and so on, which reveal the nature of rewards to a good degree. We argue that the types of reward items (called reward item types in the paper) implicitly define the nature of rewards.

Fig. 1: An illustration of the backing process on reward-based crowdfunding platforms.
Fig. 2: Comparison of projects with and without limited rewards.

Fig. 2: Comparison of projects with and without limited rewards.
Therefore, we manually examine a very large number of rewards to categorize their reward item types into fifteen types and use these reward item types to describe the implicit nature of rewards. We perform the task of automatically annotating rewards by reward item types using fifteen binary classifiers, one for each type. Finally, to verify our argument that rewards are important for project success, we evaluate how features extracted from rewards contribute to predicting the success of projects by comparing the predictiveness on project success in classifiers built using feature sets with and without reward-related features.

Contributions of this work are summarized as follows.

1) We argue for the important role of rewards played in project success in reward-based crowdfunding platforms. To the best knowledge of the authors, this is the first research on analysis of rewards in crowdfunding projects.

2) We conduct a series of statistical analyses on some explicit features of rewards to reveal their implications on project success. It is interesting to find that contradicting to the choice overload hypothesis in Marketing, the more choices of rewards offered in a project, the more likely the project succeeds. We also find that projects with limited and late-added rewards tend to receive more fundings. Furthermore, we discover that the distributions of price-goal ratios can better differentiate successful projects from unsuccessful projects, compared with the distributions of raw prices. Finally, through the Kaplan-Meier statistical analysis method, we are able to estimate the likelihood of selling out at any given time point for different price range of a reward.

3) We propose a categorization of reward item types and train classifiers for those reward item types based on reward descriptions. By annotating rewards with the reward item types, we are able to capture the nature of rewards, and facilitate further analysis on rewards in each reward item types to better understand their relationships to project success.

4) We conduct experiments to verify importance of rewards on project success. Our results show that features extracted from rewards can significantly improve the predictiveness of project success.

The rest of this paper is organized as follows. In Section II, we survey related works on crowdfunding and rewards. In Section III, we analyze the project and reward data collected from Kickstarter and discuss our findings. In Section IV, we categorize reward items into fifteen types and automatically annotate rewards into the proposed item types. In Section V, we empirically evaluate the features of rewards on prediction of project success, and analyze rewards of various reward item types. Finally, we conclude this work in Section VI.

II. BACKGROUND AND RELATED WORKS

In this section, we provide the background regarding crowdfunding platforms, particularly the reward-based crowdfunding platforms, and survey the related works.

A. Crowdfunding

Crowdfunding has raised $16.2 billion in 2014, $34.4 billion in 2015, and is expected to surpass such amounts in 2016.1 Nowadays, there are four major types of crowdfunding platforms, including equity-based, lending-based, donation-based, and reward-based.2 Among them, the reward-based crowdfunding platforms, considered as the most popular one in recent years, introduce various levels of rewards corresponding to different pledge amounts for the potential backers to choose from as a means of support to the projects.

Among the reward-based crowdfunding platforms, Kickstarter is considered as the representative one. On Kickstarter, creators offer backers with a list of rewards as options to back the projects. How the rewards are designed can potentially affect the decision made by a backer. On Kickstarter, several factors of rewards are worth looking into, including the number of rewards a project offers, the pricing of rewards, the available offerings of rewards, and the items included in each reward. However, little research has been conducted to understand how rewards affect the final results.

Researchers have been studying Kickstarter projects from various angles. One of the most widely studied research questions is: how do different factors affect the success of crowdfunding projects? Xu et al. study how the project updates influence project success [6]. Solomon et al. investigate how the timing of donations affect the final pledging results on Kickstarter [7]. Lu et al. examine how the promotions on social media affect project success [8]. Mitra et al. consider the language used to describe projects, and how they influence the fundraising results [5].

Some existing works are done on recommendation system for crowdfunding platforms. An et al. develop a system for recommending (a) the backers to projects on Kickstarter, and (b) the other way around, by considering the founders’ skill sets, pledging behavior, and projects’ geography, growing speed and categories [11]. Rakesh et al. also design a recommender for Kickstarter, by exploring features in temporal traits, personal traits, geo-location traits, and network traits [3].

B. Economics, Marketing, and Psychology View on Rewards

While various aspects of reward-based crowdfunding have been studied, we find that a study on the essential aspect of rewards is still missing. As the concept of rewards on reward-based crowdfunding platforms is similar to the items offered on a menu in places such as restaurants, we survey the literature related to the design of menu pricing to find correspondence to our study. The effect behind is well examined in the field of Psychology, e.g., from the providers’ perspective, how different strategies of pricing design can be explored on menus to maximize profits? From the consumers’ perspective, given a menu, what is a good overall design of the menu that makes the purchase most desirable [12], [13]? Regarding to the number of items provided, is there a "sweet spot" in the number of items that lead to the best results? Is it possible for a merchant to offer "too much" that leads to the phenomenon of choice overload [14]? When considering the prices, what is a good price distribution for a menu [15]?

Other than the overall design, among the menus, what characteristics of an item makes it stand out? Are rewards with limited offerings more tempting than others [16]? Are rewards with discounts more worthy than others without discounts [17]? The above questions are worth looking into.

1http://dazzinfo.com/2016/01/12/crowdfunding-industry-34-4-billion-surpass-vc-2016/
2https://www.venturebeat.com/2016/01/10/trends-in-crowdfunding/


TABLE I: Dataset statistics of the Kickstarter dataset.

<table>
<thead>
<tr>
<th></th>
<th># of projects</th>
<th># of rewards</th>
<th>Amount raised</th>
<th># of backers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Successful</td>
<td>1,399</td>
<td>15,472</td>
<td>$27,534,852</td>
<td>180,418</td>
</tr>
<tr>
<td>Unsuccessful</td>
<td>1,705</td>
<td>14,866</td>
<td>$3,758,628</td>
<td>39,483</td>
</tr>
<tr>
<td>Total</td>
<td>3,104</td>
<td>30,338</td>
<td>$31,293,480</td>
<td>219,901</td>
</tr>
</tbody>
</table>

Fig. 3: Distributions and Q-Q plot of number of rewards per project versus success rate.

III. REWARD DATA ANALYSIS

In this section, we introduce the collected Kickstarter dataset, and analyze project and reward data on Kickstarter.

A. Data Source

Kickstarter is currently the most popular reward-based crowdfunding platform. Creators with creative ideas launch projects on the Kickstarter website by describing the projects, setting pledging goals, and designing the rewards that backers receive in return if the projects succeed. Potential backers then decide whether to support the projects by evaluating the project descriptions and the rewards offered. Such a mechanism has shown to be effective in collecting funds. Since founded in April, 2009, Kickstarter has helped 103K project creators pledge over $2.29 billion. 10 million backers have been involved, and 29 million backings have been made. Being such an enormous platform for brewing creative ideas, we choose it as the source of our data analysis.

We collect data from December 15th, 2013 to March 23rd, 2014. Then we process the data to remove noises by eliminating the projects canceled, suspended, or unfinished when we end the data collection. The statistics of the final dataset are summarized in Table I.

B. Explicit Reward Features

Here we present our statistical analyses on some explicit features of project rewards on Kickstarter. These features include the number of rewards that a project offers, the prices of rewards, the limited offerings of rewards, and the later added rewards in projects’ pledging time. Note that these analyzed reward features are later evaluated by classifiers learned for predicting project success, and thus denoted as RE (Reward Explicit).

1) Number of rewards: We first observe how the number of rewards offered in a project correlates with its success rate. Intuitively, one may think that the more rewards being offered, the harder it is for the backers to choose from. This is known as the choice overload hypothesis [14]. In Psychology, how the number of options offered to a person affects their decision making is studied [18], [19]. In Marketing, some believe that a large number of options tends to make the purchase decisions of products harder for the consumers, and therefore leading to less desire to purchase the product [20]. Scheibehenne et al. state that “extensive assortments include a decrease in the motivation to choose, to commit to a choice, or to make any choice at all [14].”

In the dataset, the number of rewards offered by the projects ranges from 2 to 81, with an average of 11 rewards per project. The distribution of the number of rewards for successful and unsuccessful projects, along with the quantile-quantile (q-q) plot for the two distributions are shown in Figure 3. Notice that the q-q plot compares the distributions of number of rewards for successful projects and unsuccessful projects. The result shows that the successful projects offer significantly more rewards than the unsuccessful projects do. As shown, there are very few projects with more than 30 rewards. Therefore, we discard the projects with more than 30 rewards because they are not statistically significant. Figure 4 plots the success rate corresponding to the number of rewards ranging from 1 to 30. As shown, the higher the number of rewards offered by a project, the more likely the project succeeds. The Spearman correlation between the number of rewards offered and the project success rate is 0.964. According to our analysis, users’ backing behaviors on reward-based crowdfunding platforms go against the choice overload hypothesis.

Such a finding implies that on Kickstarter, offering more rewards is likely to satisfy a variety of potential backers with more options. In other words, having more options to choose from may raise the tendency and willingness of potential backers to support the projects. In Marketing, some believe that a large variety of choices increases the likelihood of satisfying the diverse consumers and thus caters to individuality and pluralism [21]. We exploit this finding and use the number of rewards a project offers as one of the reward features affecting project success.

\[ RE_{\text{rewards}} = \text{number of rewards} \]  \hspace{1cm} (1)

2) Pricing: Most of the time, backers decide whether to back a project not only based on whether they are interested in the project, but also whether they are willing to pay for the project with the provided reward options. Gerber et al. and Mollic claim that backers on crowdfunding platforms primarily decide their backings based on the different sizes of funding rewards for the projects [22], [23]. We therefore examine the potential relationships between rewards’ prices and the number of backings they receive. In our dataset, the minimum cost of a reward is $1, and the maximum is $10,000, with an average of $390, and 75% of the rewards fall between the prices $1 and $390. However, the distributions of reward prices are extremely right-skewed for both successful and unsuccessful projects. Thus, instead of directly analyzing the prices, we approach our analysis on prices from the following two perspectives: (1) price range, and (2) price-goal (p/g) ratio.

When considering the price range of rewards, we group rewards based on their prices into low price, middle price, and high price. To find the best price range, we divide the prices into the 1st quantile, 2nd quantile, and 3rd quantile.

TABLE II: Price range criteria for rewards.

<table>
<thead>
<tr>
<th>Raw price</th>
<th>1st quantile</th>
<th>3rd quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>$10</td>
<td>$500</td>
</tr>
<tr>
<td>middle</td>
<td>$200</td>
<td>$1000</td>
</tr>
<tr>
<td>high</td>
<td>$3000</td>
<td>$20000</td>
</tr>
</tbody>
</table>

\[ \text{Price-goal ratio} = \frac{\text{price}}{\text{goal}} \]

The Kolmogorov-Smirnov D test statistics value is 0.20261, with a p-value of 2.2e-16.
If rewards are limited, they might run out during the pledging. We hence are interested in finding how many days it will take for a reward to sell out after it is offered. Among the limited rewards, only 14.20% of them sell out, which take an average of 13.37 days. The minimum days a reward takes to sell out is one day, while the maximum days is 60 days. Figure 6 shows the histogram of the amount of time taken for rewards to sell out, relative to the pledging time of their corresponding projects. Note that we normalize the days to sell out because the pledging durations of projects on Kickstarter vary. We find that on average the rewards that sell out take 42.4% of the project’s duration time. Also, interestingly, we find a U-shape distribution, indicating that most sold-out rewards sell out at the beginning or at the very end of the pledging.

Based on the above observations, we further analyze the potential effect of reward prices on sales of limited rewards. We conduct a Kaplan-Meier survival test on the limited rewards. Kaplan-Meier is a non-parametric statistical test often used to estimate the likelihood of patients’ living after a certain treatment in the medical field [26], [27]. In medical research, the Kaplan-Meier metric, \( S(t) \), of an observed patient at time \( t \), based on the number of patients at risk \( n \) and the number of deaths \( d \), is measured as follows.

\[
S(t) = \prod_{i \in \{t_i < t\}} \frac{n_i - d_i}{n_i} \quad (4)
\]

where \( n_i \) is the number of patients at risks before time \( t_i \), and \( d_i \) is the number of deaths at time \( t_i \) in the medical applications. To apply this metric onto our application, \( n \) is the number of rewards that have not sold out, and \( d \) is the number of rewards that have sold out. We use Kaplan-Meier to assess the relationship between the rewards’ selling-out rates and the rewards’ price ranges (Figure 7). In the test, the rewards are grouped by project success and price ranges.

We find that regardless of the project success, the reward sell-out rate is inversely associated with the price of rewards. That is, a higher price range is associated with a slower sell-out rate of the rewards. We therefore extract a feature based on the reward price range inferred from the raw price, and prediction time, which signals the probability of limited rewards to sell out at a certain time. Given a project, let \( L \) be the set of limited rewards in the project. This feature is derived as Equation (5).

\[
RE_{sell-out}(t) = 1 - \frac{1}{|L|} \sum_{l \in L} S(\rho(l), t) = 1 - \frac{1}{|L|} \sum_{l \in L} \prod_{t_i < t} \frac{n_i^{p(l)} - d_i^{p(l)}}{n_i^{p(l)}} \quad (5)
\]

where \( t_i \) means all the time prior to \( t \), \( \rho(l) \) means the price range of reward \( l \), \( n_i^{p(l)} \) is the number of rewards in price range \( \rho(l) \) that have not sold out at time \( t_i \), and \( d_i^{p(l)} \) is the number of rewards that have sold out at time \( t_i \). The feature therefore becomes a function of time.
4) Late Added Rewards: According to Kickstarter, while a project is alive, the creators can add new rewards any time during the pledging. An example of a late-added reward from The HERO Belt Project is shown in Reward example 1.

In our dataset, 2.9% of all the rewards are late-added to the pledging, and 11.5% of projects have late-added rewards. The comparison between projects with and without late-added rewards is shown in Figure 8. We find that, similar to designing rewards with limited offerings, both successful and unsuccessful projects benefit from the late-added rewards. Especially for successful projects, projects with late-added rewards on average receive 1.4 times more of raised-goal ratios than those without. We therefore extract feature based on whether the project includes late-added rewards.

\[ RE_{\text{late-added}} = \begin{cases} 
1, & \text{if project has late-added reward(s)} \\
0, & \text{otherwise} 
\end{cases} \]  

IV. REWARD ITEM TYPES

In addition to the explicit features discussed in the previous section, we also study the implicit features of rewards derived from the reward descriptions. As mentioned earlier, we argue that the types of reward items (reward item types) reveal the nature of rewards in a good degree. Therefore, in Section IV-A, we categorize reward items into fifteen types and use these reward item types to describe the implicit nature of rewards. In Section IV-B, we automatically annotate rewards by the fifteen reward item types using binary classifiers.

A. Reward Item Type Categorization

On Kickstarter, a project’s rewards usually include multiple items. For example, a reward from a film project offers a t-shirt printed with the crew’s photo, and a DVD of the final film; a clothing line project offers a thank you card, designed skirt, and an invitation to the launching party. Both rewards offer the actual products of the projects (DVD and designed skirt), while the film reward offers a spin-off (t-shirt), and the clothing reward offers a thank you card, and an experience (e.g., launching party). However, instead of listing these items explicitly, they are written in reward descriptions as unstructured texts. We therefore aim to annotate rewards by the nature (types) of the items offered in them.

To annotate the rewards, we categorize the items into fifteen reward item types by manually examining a massive number of project rewards. The reward item types are introduced in Table III. Note that we do not include limited and late-added since they have been already included as explicit features. In addition, rewards on Kickstarter are incremental in prices, which encourages creators to design the rewards by adding extra or more valuable items upon cheaper rewards. Such a property is reflected by the type include the previous.

B. Reward Annotation

To automatically annotate rewards by reward item types, we develop fifteen binary classifiers, one for each item type. We perform several Natural Language Processing (NLP) tasks on the collected rewards. First, we preprocess the reward descriptions by removing unwanted punctuation, converting all characters into lowercase, and stemming all the words with Porter Stemmer [28]. Next, we construct a \(|R| \times |W|\) matrix \(X\) of ngrams with tf-idf values, where \(R\) is the set of all rewards, and \(W\) is the set of all considered ngrams, \(n = 1 \ldots 4\), and the document frequency of each ngram is over 10. To develop binary classifiers for reward item type annotation, we prepare a dataset for training and testing by randomly selecting projects to manually label their rewards with the reward item types. In the labeling process, we also remove projects not written in English. Reward examples 2 and 3 illustrate the classification of rewards into reward item types. For clarity, items in the
examples are superscripted with type number shown in Table III.

As mentioned earlier, include the previous rewards include all of the items mentioned in lower-priced rewards within the same project. We identify this item type by scanning rewards within a project after determining each reward’s item types.

In the next section, we evaluate the results of annotating rewards into the proposed reward item types (Section V-A), and analyze the rewards in each reward item type (Section V-C).

V. EVALUATION AND ANALYSIS

To verify our argument that reward information is important for reward-based crowdfunding platforms, we design experiments to demonstrate the impact of reward information (i.e., features discussed in the paper) on prediction of project success. In the following, we first evaluate the effectiveness of the developed classifiers for reward item type annotation, and then show the experimental results for evaluating reward features in project success prediction. Finally, we analyze the rewards in reward item types.

A. Reward Item Type Annotation

We first construct the ngram features of rewards and then use them to train fifteen binary classifiers (one for each reward item type) using Support Vector Machine (SVM) [29]. We train and test each SVM binary classifier using 10-fold cross validation. Notice that instead of using the same features for all the classifiers, we use recursive feature elimination (RFE) algorithm to select the top 50 features for each classifier, in order to create more concise feature sets. With this approach, we can reach an average accuracy of 97.58%, precision of 94.34%, and recall of 89.63%.

To showcase the learned classifiers, we present the ngram features of four classifiers, including thank you note, spin-offs, digital copy of products, and include the previous, in form of word clouds in Figure 9. The font sizes are scaled based on the importance learned by each classifier. As shown in the word clouds, the important words grasped by each classifier reflect the nature of the rewards. For example, rewards of spin-offs type can be detected by phrases such as shirt, sticker, and autograph.

B. Project Success Prediction

To see the value of rewards on reward-based crowdfunding platforms, we design experiments to demonstrate the importance of reward-based features in project success prediction. We formulate the problem of project success prediction as a classification problem with binary results: successful (achieving pledging goal) and unsuccessful (not achieving pledging goal). We evaluate the effectiveness of reward-related features by using project-based (i.e., non-reward) features as baseline (denoted by P). We then augment the project-based features with reward-related features, including reward explicit features (denoted by P + RE), reward implicit features (denoted by P + RI), and both (denoted by P + RE + RI), for comparison. We use four metrics to measure the performance: accuracy, precision, recall, and AUC (area under the receiver operating characteristic [ROC] curve).

The results of project success prediction are shown in Figure 10 and Figure 11. Both P + RE and P + RI outperform P, while P + RE is more effective than P + RI. Moreover, P + RE + RI achieves the best performance, leading to a 82.40% accuracy and an AUC of 0.881. This shows that our proposed reward-based features provide valuable information to improve the prediction of project success.

As discussed previously, feature REsell_out is a function of the time in project pledging duration. Thus we conduct the prediction in different stages of the project pledging, using RE and RE + RI, to see how the accuracy of prediction changes over time, as shown in Figure 12. Note that we do not include RI in this test because it does not include time-dependent features. For both feature sets, the accuracy reach the highest around half-way into the pledging duration (50% for RE and 60% for RE + RI).

We also examine the importance of features by using feature selection algorithm ReliefF [30]. Table IV shows the top ten features and the feature sets they belong to along with their weights. As shown, among the ten features, six are reward-related features, with four REs and two RIs. Also, RE and RI are measured as more important than P within the top ten features. In general, RE features are more decisive than RI,
with the most important feature being the % of rewards with high p/g ratio ($RE_{\text{high}}$), third most important being whether a project has late-added reward(s) ($RE_{\text{late-added}}$), and fifth most important as the number of rewards a project offers ($RE_{\text{rewards}}$). However, intertwined with $RE$, some extracted $RI$ features are also shown as decisive: number of physical copy of products rewards and number of thank you (social media) rewards. This exhibits the importance of both $RE$ and $RI$ features.

To further assess the feature importance within the same feature set, we also conduct ReliefF using purely $RE$ and $RI$. We find that when considering only $RE$, $RE_{\text{high}}$ is still the most important feature, while $RE_{\text{limited}}$ becomes more important than $RE_{\text{late-added}}$. Also, when considering only $RI$, we find that instead of physical copy of products, early access is evaluated as the most important, which is also shown later in Section V-C as the least popular reward item type.

### C. Analysis of Reward Item Types

In this section, we analyze the rewards in reward item types to discuss how they correlate to other aspects of Kickstarter projects, specifically the project categories, success rates, and the reward price.

To understand what kinds of rewards backers prefer, we investigate the average number of backers received and project success rates for each reward item type. Figure 13(a) shows that include the previous rewards receive on average the most backers (an average of 358.94 backers, with 3.28 coefficient of variation). This is also reflected on the project success rates. Projects with include the previous rewards have the highest average success rate (93.53%), as shown in Figure 13(b), which indicates that the backers may be tempted by rewards that include items from lower tiers of rewards. Projects with thank you rewards have a success rate of 60.01%. Even though backers do not receive any actual gift, it is possible that many backers select these rewards, leading to higher amount of funding received. It is interesting to see that spin-off rewards lead to higher success rate than digital and physical copy of product rewards. The reason behind may be due to the lower prices and small gifts. Surprisingly, projects that offer early bird and discount rewards do not lead to high success rates (50.00% and 43.59%, respectively). A possible cause of this could be the small number of offerings these rewards have. However, such a phenomenon might vary by project categories.

Next, we analyze how different categories of projects correlate with their successes for each reward item type. On Kickstarter, there are fifteen major project categories, e.g., art, music, technology, and food. To demonstrate how adoptions of reward item types differ between project categories, we plot two distributions of project categories in radar charts, as shown in Figure 14. One can see that reward item types are used very differently by the different categories of projects: spin-offs are largely adopted by film & video and comics projects, while credits are mostly adopted by film & video and design projects. Figure 15 shows the success rates of all the reward item types in each project category in form of heat map, where the darker is a block, the higher is a success rate. As shown, even though physical copy of products is highly adopted among different categories of projects, it does not correlate with high success rate for any of these project categories. On the other hand, even though experience rewards do not correlate with high success rate for most of the project categories, they tend to have a high success rate in technology projects. For instance, project Napwell: The World’s First Napping Mask not only designs rewards with different packages of napping mask products, but also offers three rewards that invite backers to have dinner with the creator, costing $200 to $400. These rewards turn
out to be very popular, leading the projects to a successful $1,564 funding, even though the original goal was $30,000. Also, we find that within a project category, different reward item types can potentially influence different levels of success rates. For example, games projects have the highest success rates when including rewards with thank you (social media). This may be due to the fact that many backers with gaming interests have a lot of online social activities. For them, having their names announced on social media is highly rewarding. For example, project Wyrd Con V originally planned to raise $8,888, while they collected $19,295 at the end. The rewards they designed heavily rely on showing appreciation to backers on social media.

VI. CONCLUSION AND FUTURE WORK

In this paper, we argue for the importance of rewards on reward-based crowdfunding platforms. We perform statistical data analysis on Kickstarter data to better understand the importance of rewards to project success. We categorize reward item types, and automatically annotate rewards with reward item types using binary classifiers. To verify our argument that reward information help predicting project success, we conduct experiments on the task using different feature sets. Our results show that augmenting both the explicit and implicit reward-based features with project-based features, we can predict the project success most accurately. Finally, we leverage the learned reward item types to analyze potential correlations between reward item types, project categories, and success of projects on crowdfunding platforms.

For the future work, we aim to explore the contents of rewards with finer granularity, e.g., extracting the semantics and similarities between rewards. We also plan to tie the analyses on rewards and projects with information on backers, in order to study the backers’ preferences not only on projects, but also on rewards. Moreover, we plan to study the rewards across different platforms to see the different effect they potentially have.

REFERENCES