Toward Automatic Bootstrapping of Online Communities Using Decision-theoretic Optimization

Shih-Wen Huang*
University of Washington
Seattle, WA, USA
wenhuang@cs.washington.edu

Jonathan Bragg*
University of Washington
Seattle, WA, USA
jbragg@cs.washington.edu

Isaac Cowhey
Allen Institute for AI
Seattle, WA, USA
isaacc@allenai.org

Oren Etzioni
Allen Institute for AI
Seattle, WA, USA
orene@allenai.org

Daniel S. Weld
University of Washington
Seattle, WA, USA
weld@cs.washington.edu

ABSTRACT

Successful online communities (e.g., Wikipedia, Yelp, and StackOverflow) can produce valuable content. However, many communities fail in their initial stages. Starting an online community is challenging because there is not enough content to attract a critical mass of active members. This paper examines methods for addressing this cold-start problem in *datamining-bootstrappable communities* by attracting non-members to contribute to the community. We make four contributions: 1) we characterize a set of communities that are "datamining-bootstrappable" and define the bootstrapping problem in terms of decision-theoretic optimization, 2) we estimate the model parameters in a case study involving the *Open AI Resources* website, 3) we demonstrate that non-members' predicted interest levels and request design are important features that can significantly affect the contribution rate, and 4) we ran a simulation experiment using data generated with the learned parameters and show that our decision-theoretic optimization algorithm can generate as much community utility when bootstrapping the community as our strongest baseline while issuing only 55% as many contribution requests.

Author Keywords

Online communities; decision theory; optimization; crowdsourcing.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI) : Miscellaneous

*Shih-Wen Huang and Jonathan Bragg contributed equally to this work.

CSCW '16, February 27-March 02, 2016, San Francisco, CA, USA

ACM 978-1-4503-3592-8/16/02$15.00
DOI: http://dx.doi.org/10.1145/2818048.2820024

INTRODUCTION

The Internet has spawned communities that create extraordinary resources. For example, Wikipedia's 23 million users have created over 4 million English articles, a resource over 100 times larger than any other encyclopedia. Similarly, StackOverflow has become a top resource for programmers with 14 million answers to 8.5 million questions, while Yelp users generated more than 67 million reviews.[1]

In reality, however, most online communities fail. For example, thousands of open source projects have been created on SourceForge, but only 10% have three or more members [24]. Furthermore, more than 50% of email-based groups received no messages during a four-month study period [6]. Since network effects sustain successful communities, the key challenge for designers is to kindle initial community activity that leads to a tipping point [13, Section 17.3].

Previous research has focused on methods for encouraging existing community members to contribute additional content. For example, SuggestBot used the edit history of Wikipedia editors to recommend articles for them to edit [9]. Beenen et al. conducted an experiment on MovieLens [10] and showed that designing requests based on social psychology theories can better motivate users to contribute [3]. Burke et al. [5] show that the community can encourage the contribution from the newcomer by showing the contribution of their friends, While these results provide insights on how to cause existing community members to increase their activity, they do not address the community "cold-start" problem. Without enough user-contributed content to attract a critical mass, there might never be enough value to recruit initial members to join the community [24].

This paper examines methods for solving the cold start problem by bootstrapping community content from the contributions of non-members. Several challenges make this a difficult problem. First, since the community doesn't have an activity log for non-members, it is hard to model their interests and recommend tasks accordingly. Second, non-members have no existing commitment to the community, so

[1] Statistics gathered in December, 2014.

they might not be inclined to make a contribution; it is unclear which social psychology theory one could use to encourage contributions in this case. Finally, there are a huge number of possible non-members and many candidate tasks to suggest; determining which requests should be sent to which users represents a combinatorial optimization problem.

Specifically, we identify a class of communities, which we call *datamining bootstrappable*, where an external resource provides a means of identifying potential members and estimating their interests and expertise. For these communities, we define the bootstrapping process as a decision-theoretic optimization problem. Previous research has shown decision-theoretic optimization is useful in similar social computing contexts such as crowdsourcing [11]. Applying the decision-theoretic framework to model the bootstrapping problem allows us to estimate the utility of different operations and find a set of operations that are near optimal for the community.

In addition, we conducted a field experiment on *Open AI Resources* (Open AIR)[2], a website which launched in July 2014, three months before our study. In collaboration with the Allen Institute of Artificial Intelligence (AI2)[3], we were allowed to access the user data of the site. This provided us a unique opportunity to study the community bootstrapping problem because Open AIR hadn't accumulated much reputation or user-generated content when we conducted our study.

By text-mining information from the Google Scholar citation graph and linking to author homepages, we identified individuals who might be willing to join the Open AIR community. We considered a range of strategies to get these individuals involved and measured their response rates. The results from these experiments inform the parameters of a decision-theoretic model that can control the community bootstrapping operation. Our study is an initial step toward building an automatic system that can bootstrap the contents of online communities. In summary, our paper makes the following contributions:

1. We characterize a class of online communities, which we call "datamining bootstrappable communities," where an external resource provides a means of identifying potential members and estimating their interests and expertise. We then define the problem of efficiently bootstrapping such a community in terms of decision-theoretic optimization, and propose a greedy algorithm that efficiently solves this problem with performance guarantees.

2. Using the *Open AI Resources* community as a case study, we identify a set of informative, text-minable features and estimate the probabilities that parametrize the actions in our model.

3. We demonstrate that request design, such as the *foot-in-the-door technique* [16], and the estimated level of interest of non-members are important features in the model that can significantly affect the probability of contribution.

4. We ran an experiment using synthetic data generated with parameters learned from the real data we collected to show that our decision-theoretic optimization algorithm can achieve comparable utility for bootstrapping online communities while issuing only 55% as many requests, compared to the strongest baseline.

[2]*Open AI Resources*, `http://airesources.org/`, is an online community that allows users to comment and discuss Artificial Intelligence (AI) related open-source datasets and software.
[3]`http://www.allenai.org/`

RELATED WORK

Starting New Online Communities

Online communities are virtual spaces where people can interact with each other. Many new online communities fail because they are unable to carve out a useful niche, to provide enough value to accrete a community, or because they lose to competition from other communities [24].

There are several ways to help a community reach critical mass. One popular method is to leverage existing members to recruit new members. Previous research has shown that a person is more likely to join a community if he or she has friends that are already members [2]. Companies, such as Dropbox, exploit this principle by providing incentives for users to refer their friends [24].

Bootstrapping the content of online communities is a complementary approach to the cold-start problem and especially useful when resulting content is long lived. Seeded content can increase the utility for initial members to join the community [24]. Therefore, many online communities bootstrap by copying content from 3rd parties. For example, MovieLens imported a database of ratings from another movie rating website (EachMovie.com), which was no longer operational. Resnick et al. [28] show that using paid staff to prepopulate the forum made it more attractive for other people to post to and read the board. Seeded content not only increases the utility of users, encouraging them to join the community, but also can be used to direct the behavior of the new users, encouraging them to contribute similar content [32].

Intelligent Task Routing in Online Communities

To make an online community thrive, the community designer needs to find ways to encourage contributions from its members. One approach, called *intelligent task routing*, models each member's interests, determines a task of potential interest, and sends them a personalized request [8]. As one example, Cosley et al. [9] utilized the edit history of Wikipedia users to model their interests. Their system used information retrieval and collaborative filtering techniques to suggest tasks, significantly increasing the contribution rate. Another approach utilized the rating history of MovieLens users to find those whose ratings differ the most. Their system used this information to send personalized suggestions encouraging users to reply to forum posts of users with opposite opinions; this significantly increased the reply rate [20].

These exciting results show that intelligent task routing can be used to encourage community members to contribute additional content. However, it is not clear whether one can use intelligent task routing to bootstrap content from non-members, who have not generated activity logs that help with targeting.

Encouraging Contribution Using Request Design

Research in social psychology has shown social influence or persuasive techniques can be used to make people more likely to comply with requests [7]. Since the success of online communities relies heavily on the contributions of community members, many studies have been done to examine how request design can be used to encourage member contributions [24]. For instance, Burke et al. [4] show that asking more specific questions can increase the response rate by 50%. Also, using the social psychology theory of social loafing, Beenen et al. [3] show that requests stressing the uniqueness of the member's contribution significantly increased the contribution rate. Moreover, Lopez and Brusilovsky [25] suggest that the system should adapt the design based on user demographics. All these studies focus on designing requests to get contributions from community members. Users that have not yet committed to the community might be less likely to accept the requests because they are less invested in the community. Therefore, we might need to find another theory in social psychology to motivate request designs that are more suitable for encouraging contributions from non-members.

Research has shown that people are more likely to respond to a large request after they have accepted a smaller request because they want to maintain the consistency of their self-perception [7]. Therefore, one compliance technique that is often used in industry is to hide the large request and present the smaller request to the recipients first, a method called the "foot-in-the-door technique" [16]. Gueguen [19] showed that this method is not only useful in a face-to-face scenario, but can also be used in computer-mediated communication (e.g., email). However, to the best of our knowledge, no researchers have investigated whether an online community could use the foot-in-the-door technique to bootstrap contributions.

DECISION-THEORETIC COMMUNITY BOOTSTRAPPING

Background

Numerous researchers have applied decision theory to control interface behavior. For example, the BUSYBODY system mediates incoming notifications using a decision-theoretic model of the expected cost of an interruption, which is computed in terms of the user's activity history, location, time of day, number of companions, and conversational status [21]. LINEDRIVE illustrates driving directions using optimization to find the optimal balance between readability and fidelity to the original shapes and lengths of road segments [1]. SUPPLE [17] renders interfaces using decision theory to optimize the ease of a user's expected behavior. Researchers have also used decision theory to control social interactions, especially for the specific application of optimizing crowdsourced workflows [11, 23, 33]. Our model of community bootstrapping, described in the next section, builds on this seminal work.

Applicable Communities

We start by precisely a class of online communities where our bootstrapping methods are appropriate and formally define the problem of eliciting contributions. Since the very notion of community is amorphous, we assume that there are a set of humans $\mathcal{H}$ who are potentially willing to make contributions C to different tasks $\mathcal{T}$. For example, for Yelp, $\mathcal{T}$ might be the set of restaurants, and C could be a contributed review. For AirBnB, there might be a single task (list a house) and each contribution would correspond to a rental property. We note that many communities have complex (two-sided) market dynamics [29], which we are ignoring; however, from a practical point of view, a single side tends to dominate most markets. For example, AirBnB focused on sellers, since their contributions (available rental inventory) were durable (led to repeated transactions); renters came easily. Similarly, Wikipedia was bootstrapped with an early focus on authors, even though the community would fail without readers as well.

We say a community is potentially *datamining-bootstrappable* if there are mechanisms, likely using external websites or similar resources, for satisfying the following conditions:

1. Identifying the humans who are potentially interested in a given task, and estimating the probability that they will contribute.
2. Sending a request to a those humans (e.g., via a data-mined email address or other communication channel)
3. Estimating the quality of their contribution.

If these conditions hold, our method is applicable. However, we caution that these conditions don't guarantee that our method will bring the community to the tipping point. This depends on the specific parameters, e.g., number of humans, response rate, and actual utility of contributions, including their durability.

Request Model

Since we are targeting non-members who are not committed to the community, we assume that the community can send, at most, one contribution request to each human. In addition, since the request design greatly affects the response rate, a system can explore different requests in the design space $\mathcal{D}$. Thus the space of possible requests is $\mathcal{R} = \mathcal{H} \times \mathcal{T} \times \mathcal{D}$. Should a request result in a contribution, the quality of that contribution will come from a set of possible qualities Q. Our objective is to issue the set of requests $\mathbf{R} \subseteq \mathcal{R}$ with maximal expected utility, while satisfying the constraint that no human is asked to do more than one task. Letting $\mathbf{R_h} \subseteq \mathbf{R}$ denote the subset of requests given to human h, this constraint requires $\forall h, |\mathbf{R_h}| \leq 1$.

Probability and Quality of Contributions

The expected utility of a set of requests is defined in terms of the probability of the humans responding to the appeal and the utility of the resulting contributions. As discussed in the background section, previous work has shown that the probability of a request being honored is a function of two key factors: the human's preexisting interest in task t, which we denote $i_{h,t}$, and the request design d [3, 8, 9]. Therefore, we model the probability that human h will contribute to task t as $P(c_{h,t} | \ d, \ i_{h,t})$. Similarly, we condition the quality of a contribution on the human and their interest in the task: $P(q_{h,t} | \ h, i_{h,t})$.

In order to apply our model to a specific domain, one needs to specify a set of designs and a way of estimating interest

levels and then measure the conditional probabilities. In the Experiments section we show how this may be done for our case-study domain, the Open AIR website. For example, our experiments show that if an author has written a paper citing a resource, then this connotes a significantly higher level of interest and increased contribution rate (e.g., see Table 3).

The Utility of Contributions

In general, it is extremely difficult to estimate the quality of a given contribution [22]. As a result it is common to assume (*all other things being equal*) that more contributions are generally better than fewer. For example, Amazon emails all purchasers to solicit a review. This intuition can be formalized as monotonicity: Let A, B and N be sets of contributions. A utility function, $U : 2^N \rightarrow R$, is *monotone* if for every $A \subseteq B \subseteq N, U(A) \leq U(B)$.

Furthermore, some contributions are more valuable than others. For example, the first review of an open-source code library is probably more useful than the 100th. In general, the marginal value of a contribution to a task is smaller when the task has already received other contributions. This "diminishing returns" property is captured by the notion of submodularity. More precisely, a utility function is *submodular* if for every $A \subseteq B \subseteq N$ and $e \in N : f(A \cup \{e\}) - f(A) \geq f(B \cup \{e\}) - f(B)$ [27].

For example, one possible monotone submodular utility function can be defined in terms of the utility of the contributions to task t as $\mathrm{U}_t(\mathbf{c_t}) = \alpha \log\left[\beta \sum_{c_{h,t} \in \mathbf{c_t}} f_q(c_{h,t})\right]$, where α and β are constants, and f_q is a measure of the quality (e.g., the length) of contribution $c_{h,t}$ made to task t. But this is just an example. For the rest of the paper, we assume that the utility provided by a set of contributions to task t is a monotone submodular function U_t. For simplicity, we further assume that the system's overall utility is a linear sum of the utilities achieved on each task. This approximation is common practice [31, Section 16.4.2] and makes intuitive sense. For example, the utility Yelp receives for the reviews of restaurant A are roughly additive with those of restaurant B.[4] We seek to send out the following set of requests, which maximizes the total expected utility $\mathrm{E}[\mathrm{U}] = \mathrm{E}[\sum_{t \in \mathcal{T}} \mathrm{U}_t]$:

$$\mathbf{R} = \arg\max_{\mathbf{R} \in 2^{\mathcal{R}}} \sum_{t \in \mathcal{T}} \sum_{\mathbf{Q_t} \in 2^{\mathbf{R_t}}} P(\phi_c(\mathbf{Q_t}, \mathbf{R_t})) \sum_{\mathbf{q_t} \in |Q|^{|\mathbf{Q_t}|}} P(\phi_q(\mathbf{q_t}))\mathrm{U}_t(< \mathbf{Q_t}, \mathbf{q_t} >), \quad (1)$$

where $\mathcal{R}$ is the set of all possible requests, $\mathbf{R_t} \subseteq \mathbf{R}$ is the set of requests for task t, $\phi_c(\mathbf{Q_t}, \mathbf{R_t})$ denotes the event where $\mathbf{R_t}$ requests are made but only $\mathbf{Q_t}$ actually result in contributions, and $\phi_q(\mathbf{q_t})$ denotes the event where each of the requests in $\mathbf{Q_t}$ results in a contribution with a quality from Q, the set of possible contribution qualities. The probability of these events are defined as

$$P(\phi_c(\mathbf{Q_t}, \mathbf{R_t})) = \prod_{<h,t,d> \in \mathbf{Q_t}} P(c_{h,t}|\ d,\ i_{h,t}) \prod_{<h,t,d> \in (\mathbf{R_t} - \mathbf{Q_t})} 1 - P(c_{h,t}|\ d,\ i_{h,t})$$

and

$$P(\phi_q(\mathbf{q_t})) = \prod_{<h,t,d> \in \mathbf{Q_t}} P(q_{h,t} \mid h, i_{h,t}).$$

Recall also that $\mathbf{R}$ must satisfy the constraint that $\forall h, |\mathbf{R}_h| \leq 1$.

[4] One might argue that diminishing returns might apply *across tasks* as well as within tasks; we hope to consider this and other elaborations in future work.

Solving the Optimization Problem

Although there are various utility functions community designers can choose from, in a realistic setting, one needs to consider that the utility of each contribution depends on the contributions from other people. For instance, the utility of a contribution to a task might decrease as the number of contributions that a task has increased. Therefore, to find the solution to this optimization problem, the system needs to enumerate all the possible allocations of the requests. However, this creates $|\mathcal{T}|^{|\mathcal{H}|}$ possible allocations. If the number of humans or tasks is reasonably large (e.g., in the hundreds), then the search space will be intractable. We must therefore consider approximate solutions.

While many methods for heuristic search have been proposed, few offer performance guarantees. However, the submodular nature of the utility function allows us to closely approximate the optimal value with the following method. Algorithm 1 first computes the expected utilities for all possible requests. Then, the system sends the request with the highest expected utility. Once the system assigns one human to a task, it adjusts the expected utilities for the requests of that task and all the unassigned humans based on the expected contributions of the task. By iterating this process until all humans in $\mathcal{H}$ are assigned, the system can make sure the community has requests with the highest expected utility at each point when a partial assignment is made.

Input : $\mathcal{H}, \mathcal{T}, P(c_{h,t}|\ d,\ i_{h,t}), P(q_{h,t}|\ h,\ i_{h,t})\ \forall < h, t, d > \in \mathcal{R}$
Output: $\mathbf{R}$
Compute the initial expected utilities for all possible requests $r_{h,t,d^*}\ \forall h \in \mathcal{H}\ \forall t \in \mathcal{T}$, where $d^* = \arg\max_d P(c_{h,t}|d, i_{h,t})$;
Initialize $\mathbf{R} \leftarrow \emptyset$;
while *There is a human* $h \in \mathcal{H}$ *with no assigned task* **do**
  Find the request r_{h^*,t^*,d^*} with highest expected utility;
  Add r_{h^*,t^*,d^*} to $\mathbf{R}$;
  Mark h^* as assigned;
  Recalculate the expected utility for the requests r_{h,t^*,d^*} for every unassigned human $h \in \mathcal{H}$;
end

Algorithm 1: A decision-theoretic approximation algorithm for issuing bootstrapping requests.

Performance Guarantee

As we now show, our assumption that U_t is monotone submodular guarantees that our solutions will be good. Since each outcome is associated with a nonnegative probability and monotone submodular functions are closed under nonnegative linear combination, the expected utility of a set of

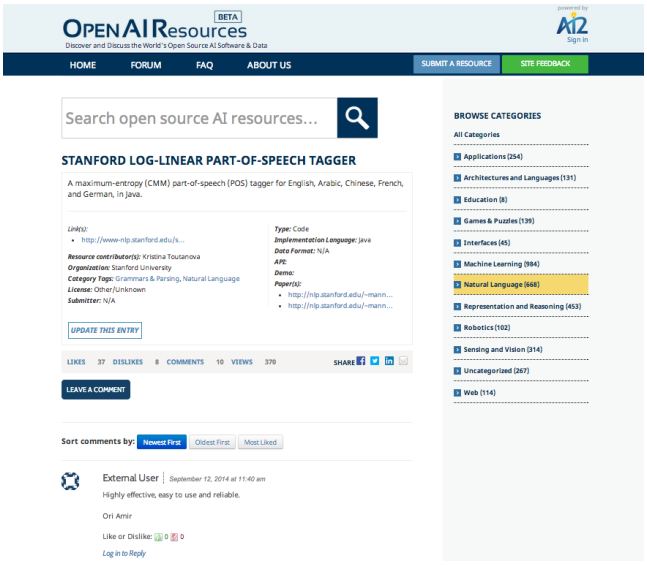

Figure 1. An example Open AI Resources (Open AIR) interface of an AI resource. The interface presents the basic information of the resource (e.g., a summary of the resource and its main contributor) and the comments of previous users. The users can update the entry or leave a comment by clicking the buttons on the interface.

requests $\mathbf{R}$ (Equation 1) is also monotone submodular.[5] This enables us to provide the following optimality guarantee for our algorithm.

THEOREM 1. *Given the constraint that at most one request may be sent to any potential contributor, Algorithm 1 obtains at least* $\frac{1}{2}$ *of the total achievable expected utility.*

PROOF. The constraint that each human receive at most one task is naturally encoded as a matroid constraint on the set of total possible requests. Fisher et al. [15] show that subject to a matroid constraint, Algorithm 1 achieves at least this fraction of the optimal value for monotone submodular functions. □

Note that we have assumed that utility functions are defined on a per-task basis, so Algorithm 1 needs only recompute at most $|\mathcal{H}|$ values in each step, leading to a worst-case $O(|\mathcal{H}|^2)$ performance.[6]

APPLYING THE MODEL TO OPEN AIR

So far, our request-assignment model has been abstract and hence applicable to many different communities. To make it concrete, we consider Open AIR (Figure 1), an online community hosted by the Allen Institute of Artificial Intelligence (AI2) which allows people to research and review open source AI resources (e.g., datasets and software). To initialize the content of Open AIR, an administrator manually searched for resources on the Web and added their information to the website. Open AIR provides three ways for a user to contribute to the community: first, by submitting a new resource; second, by updating an existing resource; third, by commenting on a resource.

We single out *commenting on a resource* for three reasons. First, the initialization process meant that Open AIR already had many resources in its DB. Second, commenting on a resource requires less effort than other kinds of contributions and should be easier to encourage. Third (and most important), reviews and opinions (comments) provide useful information for users who want to use the resources in the future, which creates unique value for the community. Bootstrapping Open AIR, therefore, means encouraging enough non-members to come and review resources that the site reaches a tipping point and becomes self-sustaining.

In order to apply the decision-theoretic model we need a set of possible contributors, $\mathcal{H}$, whom we consider to be authors with a Google Scholar page. We define one task, $t \in \mathcal{T}$, for each resource. A contribution request also has a design $d \in \mathcal{D}$, which is an email template (described below) that we use when asking the non-member to contribute a review. The final requirement for the decision-theoretic model is a set of parameter values — specifically, the probability of contribution, $P(c_{h,t}|\ d,\ i_{h,t})$, for different designs, d, and different interest levels, $i_{h,t}$. These probabilities are estimated in our Experiments section. But before we can measure these probabilities, we need to define the features that we'll use for conditioning. The next subsection discusses the set of values we consider for $i_{h,t}$ and how text mining can extract these features from public information on the Web. The following subsection describes our request designs, d.

Text Mining Features to Predict Contributions

As we have mentioned, the interests of non-members are difficult to model because the system doesn't have logs of their activity in the community. Therefore, to estimate the interest level $i_{h,t}$ of non-member h responding to a task t about a resource, we propose that a system can use text mining of publicly available information on the Web. For a research-oriented community like Open AIR, non-members of particular interest (researchers) leave information traces in the form of publications. When authors cite a resource, they indicate basic knowledge or understanding of the resource and may be more likely to write comments for that resource.

Moreover, the text surrounding a citation may contain valuable information about the citation, its role in the paper, and the author's interest in the corresponding resource. To analyze these citation contexts, we manually examined 100 of them to see if there were any patterns. Preliminary analysis revealed two types of citation contexts: 1) the authors indicate using the software or dataset to conduct an experiment; or 2) the authors do not indicate using the resource, but list the

[5]Moreover, this function is *adaptive monotone submodular* [18], since we assume contribution probabilities are independent. While this means that an adaptive version of Algorithm 1 is also near-optimal, adaptive requests are impractical due to delays between the request and contribution; we do not consider this setting further in this paper.

[6]Computing utilities involves taking an expectation over the possible outcomes for the requests assigned to a task. We assume that the probability and value of human contributions come from a bounded number of classes (Table 3), which enables speedy utility computations.

TEXT-USE	Citation Context
True	We used the Stanford Named Entity Recognizer [4] in order to extract names of places, organizations and people names from the target.
True	We apply Stanford NER toolkit to extract named entities from the texts (Finkel et al., 2005).
False	In contrast, NER systems only categorize named entities to several predefined classes (typically 'organization', 'person', 'location', 'miscellaneous' [13]).
False	However, current NER systems such as Stanford NER that achieve F1 scores of 0.87 on news articles [21], achieve a significantly lower F1 score of 0.39 on tweets with a precision as low as 0.35.

Table 1. Examples of the two types of citation context that determine the TEXT-USE feature.

paper to recognize previous or related work. For examples of the two types of citation contexts, see Table 1.

An orthogonal dimension of the citation context is the sentiment of the text. If an author expresses strong sentiment, either positive or negative, he or she may be more likely to respond to a request about the cited resource [12]. We enumerate the text mining features that characterize the interaction $i_{h,t}$ of non-member h with a task t as follows:

- **TEXT-CITE**: *True* if a human, h, has cited the resource and *False* otherwise.
- **TEXT-USE**: *True* if h has cited the resource and indicated use and *False* if h has cited the resource without indicating use.
- **TEXT-SENT**: *Positive*, *Neutral*, or *Negative* based on the sentiment of h's citation of the resource.

TEXT-USE and **TEXT-SENT** are only defined when **TEXT-CITE** is True.

For each resource, we determine these features for non-members as follows. First, we extract the title of the paper that describes the resource. For instance, the Open AIR resource "Stanford log-linear part-of-speech tagger" has a link to the associated paper "Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger." Using the title of this paper, we search Google Scholar to retrieve the papers that cite the resource paper using the "Cited by X" link. [7] Finally, we parse the authors' email addresses in the citing papers to obtain a list of emails of non-members who have written papers that cite the resource paper. **TEXT-CITE** is *True* for a non-member / resource pair when the non-member's email appears in this list.

[7]Our experiments adhere to the Google Scholar ToS, as the first author manually executed the queries.

To extract the citation contexts, we built a parser to parse the text around citations in two common reference styles. The first style is the (*Author year*) format used in this paper. To find these citations, we used a regular expression to construct the (*Author year*) pattern with the author information and published year of the resource paper. The second reference style is the [*number*] format, where the *number* is the index of the citation in the paper's References section. To find these citations, we searched for the title of the resource paper in the References section to find its index, and searched for the pattern [*number*] using that index. In each case, the citation context is the sentence containing the citation.

We analyze these citation contexts to determine the values of **TEXT-USE** and **TEXT-SENT**. To determine the value of **TEXT-USE**, we expanded the verb *use* with WordNet [26] to obtain a synset, or collection of synonym words.[8] **TEXT-USE** is *True* when any of the stemmed words in the citation context matches a word from the synset. To determine the value of **TEXT-SENT**, we perform sentiment analysis of the citation context using SentiWordNet [14]. SentiWordNet is a lexical resource that maps each synset in WordNet to scores that quantify the amount of positive and negative sentiment. To calculate a single average score for the sentiment of the citation context, we sum the positive scores and subtract the negative scores of each word, then divide by the total number of words. **TEXT-SENT** is *Negative* for average scores lower than -0.007 (1st quartile), *Positive* for scores higher than 0.018 (3rd quartile), and *Neutral* for scores between -0.007 and 0.018.

Designing Contribution Requests

Another parameter that affects the probability of contribution in our model is the request design d. In our case, this refers to the type of email message that was sent to non-members asking for a comment. When crafting these pleas, we expected that the response rate would be inversely proportional to the effort needed for the person to make the contribution, so we tried to make contributions as simple as possible.

Our initial design allowed non-members to comment on a resource by simply hitting "reply" to the request email; the body of their reply was automatically added to the Open AIR website and attributed to the person sending the email. Although this allowed people to enter reviews quickly and without leaving their email client, it was a failure. Not a single one of the 69 recipients submitted a review. Informal interviews suggested that the problem was likely a lack of context — the non-members had neither a sense of the type of review that was expected nor how it would appear on the website. Based on this feedback, we switched to a different approach.

Our next designs brought the non-member directly to the site. A baseline design explained Open AIR and how their contribution would benefit the community, then provided a hyper-link labeled "Tell us about your experience using this resource." If the person clicked on the link the landing page presented two text boxes for 1) their name and 2) their comments on the resource.

[8]The words in the synset are *use*, *utilize*, *apply*, and *employ*.

Resource Name	Associated Paper
Scikit-learn	Scikit-learn: Machine learning in Python
WEKA	The WEKA Data Mining Software: an Update
Pascal VOC Dataset	The Pascal Visual Object Classes (VOC) Challenge
Stanford Named Entity Recognizer	Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling
Caltech 101	Learning Generative Visual Models from Few Training Examples

Table 2. The five resources that we asked the non-members to comment on in our study.

Our final design utilized a method known as the "foot-in-the-door technique" [16], which showed that if one first asks a person to complete an easy task, they are more likely to later do a more time consuming task. With this method the candidate received a message identical to the baseline, but instead of a link inviting "Tell us about your experience using this resource." we presented a simple question which asked whether the recipient would like to recommend the resource to other AI researchers, and then presented two links, one for "Yes" and one for "No." Once the non-member clicked one of these links an Open AIR page would open in a web browser displaying an interface that invited them to elaborate their opinion with more detailed information In summary, we experimented with two designs:

- **REQ-FOOT**: *True* if a design, *d*, uses the *foot-in-the-door* techniques and *False* otherwise.

Although the comments induced in the **REQ-FOOT** = *True* condition require the same amount of effort as the baseline (clicking a link, entering a comment, and clicking submit), we conjectured that the the foot-in-the-door design would yield a greater number of reviews. The ability to contribute a useful bit of information with a single click might induce non-members to invest in the community, and once engaged they would more likely contribute a review.

EXPERIMENTS

To estimate the probability of contribution for different conditions, we conducted a set of controlled experiments with Open AIR. In our experiments, we focused on five different Open AIR resources (Table 2) and emailed contribution requests to 1,339 non-members who cited at least one of the resources. The emails were sent out on weekday mornings between 10/28/2014 and 11/10/2014. For reference, the email template for the request which applies the foot-in-the-door technique can be found in the Appendix. The email campaign was managed using MailChimp,[9] which allowed us to record whether the recipients opened the email and clicked the links in the email. If the non-members accepted the request and

[9] http://mailchimp.com/

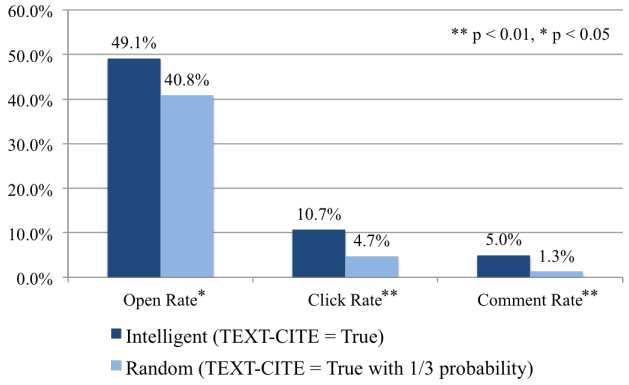

Figure 2. Comparison between the requests that perfectly match the non-member and the resource they cited (Intelligent), and the requests randomly assign one of the three resources to the non-member (Random). The results show that the requests that perfectly match the non-members and the resources they cited had a significantly higher open rate, click rate, and comment rate.

commented on the resources using the web interface, our program would record the information of the comments and automatically post the comments on Open AIR. To examine the quality of the comments generated by these non-members, we manually examined all the comments resulted from the requests of our experiments. The average length of the comments is 26.4 words. The comments also provide different perspectives for the users to better understand the resources. For example, one comment mentioned that Scikit-learn is not only a useful resource, one can also learn about the algorithms from their website and documentation: *Scikit provides a wide variety of Machine Learning and data-processing algorithms, all interfaced through Python. Plus, their website is a great resource for concepts and details about the algorithms.* Also, another comment about PASCAL VOC dataset helps the users to understand why this dataset is so important to computer vision research: *Currently, it is the best computer vision dataset for evaluating object detection algorithms. It has had a long history and has been instrumental in greatly improving the state-of-the-art of object detection*. The results indicate that the comments generated by these non-members can be really useful for other members in the community

The Effects of Citing a Resource

In the previous section, we made the case for the following hypothesis.

H1a: *Non-members who wrote papers that cited a resource are more likely to accept a request to comment on the resource.*

To see if we can use the information of who cited the resource (obtained by text mining the Web) to bootstrap contributions from non-members, we conducted an experiment that sent requests to non-members who cited one of three Open AIR resources: Weka, Scikit-learn, and the Pascal VOC Dataset. The emails were sent under two conditions:

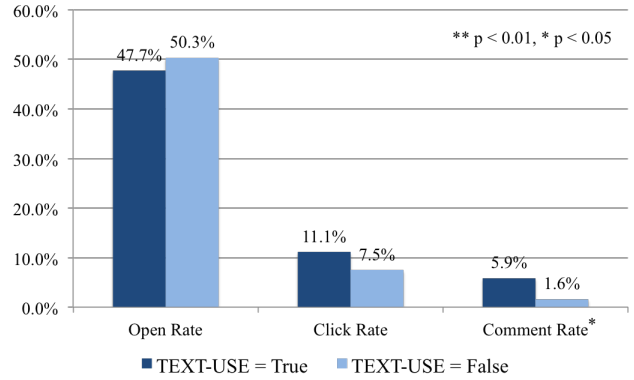

Figure 3. Comparison between the requests sent to non-members who expressed they used the resource in the citation context and those who didn't. The results show that the non-members who mentioned they used the resource in the citation context were significantly more likely to comment on the resource.

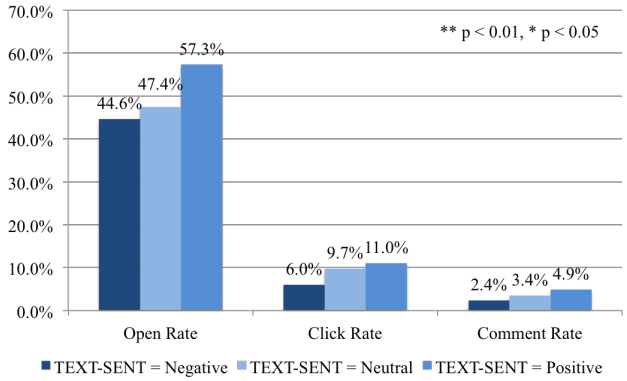

Figure 4. Comparison between the requests sent to non-members who expressed negative, neutral, and positive opinions in the citation context. The results show that there were no significant differences between the conditions.

1. **Intelligent**: The system sent contribution requests only to non-members who have cited one of the three resources (**TEXT-CITE** = *True* for all *h* in this condition).
2. **Random**: The system randomly selected one of the three resources and sent requests asking the non-members to comment on that resource. We controlled this condition to ensure that the probability a recipient has cited the resource was at least 1/3 by including the non-members that we knew cited the resource (P(**TEXT-CITE** = *True*) $\geq 1/3$ in this condition).

403 request emails were sent under the Intelligent condition and 382 request emails were sent under the Random condition. The results show that the requests sent under the Intelligent condition had significantly higher email-open rates (49.1% v.s. 40.8%, $\chi^2 = 5.45$, $p = 0.02$, $n = 885$, $df = 1$, effect size= 0.079), link-click rates (10.7% v.s. 4.7%, $\chi^2 = 9.71$, $p < 0.01$, $n = 885$, $df = 1$, effect size = 0.105), and comment rates (5.0% v.s. 1.3%, $\chi^2 = 8.49$, $p < 0.01$, $n = 885$, $df = 1$, effect size = 0.098) (Figure 2). One should note that the Random condition tested in this study is actually quite a strong baseline with 1/3 chance that the non-members cited the resource. Moreover, subsequent analysis showed that all of the non-members who ended up writing comments under the Random condition in fact *had* cited the corresponding resource in some paper. This result provides strong support for **H1a** and shows that authors that cited a resource (**TEXT-CITE** = *True*) were significantly more likely to open email requests, follow links, and contribute by writing comments about the resource. It also confirms our hypothesis that obtaining features by text mining the web can be used to help a community bootstrap content from the contributions of non-members.

The Effects of the Context of a Citation

Since the context of a citation provides information about an author's relationship to a cited resource, we are interested in whether we can use text mining to determine which authors are more likely to accept contribution requests. Our first hypothesis is that the **TEXT-USE** feature will help us predict requests that will result in contributions.

H1b: *Non-members who indicate in the citation context having used a resource are more likely to comment on the resource.*

In this experiment, we focused on 340 non-members who cited one of the five resources listed in Table 2. We extracted the context of their citation, using the method described in the previous section. We separated the non-members into two conditions, based on whether the citation context showed that they had used the resource to accomplish a task (**TEXT-USE** = *True*, $n = 153$) or merely compared to it (**TEXT-USE** = *False*, $n = 187$).

The results show that the email-open rates were similar for the two conditions (47.7% v.s. 50.3%, $\chi^2 = 0.22$, $p = 0.64$, $n = 340$, $df = 1$, effect size = 0.025). In terms of click rates, we can see that the click rate for the group who indicated using the resource is about 50% higher, but the difference is not statistically significant (11.1% v.s. 7.5%, $\chi^2 = 1.33$, $p = 0.25$, $n = 340$, $df = 1$, effect size = 0.063). The biggest difference between the two groups occurred in the comment rate. Non-members whose context indicated resource usage were three times more likely to provide a written review than the control group (5.9% v.s. 1.6%, $\chi^2 = 4.52$, $p = 0.03$, $n = 340$, $df = 1$, effect size = 0.116) (Figure 3). This supports **H1b** and shows that authors who indicate having used the resource in the citation context (**TEXT-USE** = *True*) have a significantly higher contribution rate.

Previous research has shown that people are more likely to leave highly positive or negative reviews because high valence experiences often motivate interpersonal communication [12]. Therefore, we hypothesized that this might also apply with respect to request acceptance.

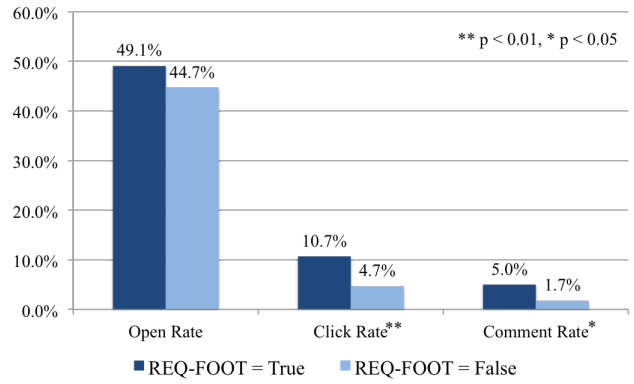

Figure 5. Comparison between the requests that asked a simple yes/no question first and the baseline. The results showed that requests that applied the foot-in-the-door technique lead to more clicks and more comments from the non-members.

H1c: *Non-members who expressed strong positive or negative opinions when describing the citation are more likely to comment on the resource.*

To test this hypothesis in the case of Open AIR, we analyzed the same 340 emails from the previous experiment, partitioning them into 3 groups: **TEXT-SENT** = *Negative* (n = 83), **TEXT-SENT** = *Neutral* (n = 175), and **TEXT-SENT** = *Positive* (n = 82), based on their average sentiment score.

The results suggest that there were no significant differences between the sentiment score groups in terms of email-open rate (44.6% v.s. 47.4% v.s. 57.3%, χ^2 = 3.09, p = 0.21, n = 340, df = 2, effect size = 0.096), link-click rate (6.0% v.s. 9.7% v.s. 11.0%, χ^2 = 1.38, p = 0.50, n = 340, df = 2, effect size = 0.064), or comment rate (2.4% v.s. 3.4% v.s. 4.9%, χ^2 = 0.75, p = 0.69, n = 340, df = 2, effect size = 0.047) (Figure 4). There are several possible reasons that we couldn't find significant differences across these conditions. First, our sample size may have been too small. Secondly, an author's sentiments may have changed since the paper was written. Furthermore, the sentiment analysis we performed was imperfect, and some contexts may have been mis-classified. We believe a follow-up study may be warranted. Nevertheless, at least in our current experiment, we were unable to find evidence for **H1c** and we conclude that a person's stated sentiment toward a resource (**TEXT-SENT**) might not be an important factor for contribution rate.

The Effects of Foot-in-the-Door Request Design

Since the foot-in-the-door technique has been proven in a business context [16], we hypothesized that it might apply to online communities.

H2: *Non-members who initially receive a smaller "Yes/no" request are more likely to subsequently contribute a written review to the community.*

To test this hypothesis, we sent 403 request emails which asked a yes/no question first (see Appendix) and then invited a written comment (**REQ-FOOT** = *True*) and 407 request emails that directly asked the recipients to write a comment (**REQ-FOOT** = *False*). Recipients in both conditions were non-members who had cited either Weka, Scikit-learn, or the Pascal VOC Dataset. The email-open rates, link-click rates, and comment rates of the two conditions were compared.

d	$i_{h,t}$	$P(c_{h,t}\|d, i_{h,t})$
REQ-FOOT	**TEXT-CITE**	0.050
REQ-FOOT	**TEXT-USE**	0.059
REQ-FOOT	**TEXT-CITE** $\wedge\neg$ **TEXT-USE**	0.016
REQ-FOOT	**TEXT-CITE** $\wedge$ **TEXT-SENT**=*Pos*	0.049
REQ-FOOT	**TEXT-CITE** $\wedge$ **TEXT-SENT**=*Neut*	0.034
REQ-FOOT	**TEXT-CITE** $\wedge$ **TEXT-SENT**=*Neg*	0.024
$\neg$ **REQ-FOOT**	**TEXT-CITE**	0.017

Table 3. The contribution probabilities with different design requests and the interest of a human in a task.

The results showed that the email-open rates were similar between the two conditions (49.1% v.s. 44.7%, χ^2 = 1.58, p = 0.21, n = 810, df = 1, effect size = 0.044), which makes sense because the subject lines were identical. However, the non-members who received foot-in-the-door requests were not only significantly more likely to click the link in the email (10.7% v.s. 4.7%, χ^2 = 10.32, p < 0.01, n = 810, df = 1, effect size = 0.113), they were also significantly more likely to leave comments about the resource (5.0% v.s. 1.7%, χ^2 = 6.61, p = 0.01, n = 810, df = 1, effect size = 0.906) (Figure 5). Thus, the findings support **H2** and show that the foot-in-the-door technique is an important tool for encouraging comments.

SIMULATION EXPERIMENT

To examine whether the decision-theoretic model we proposed really increases the utility of bootstrapping online communities, we conducted a simulation experiment using the synthetic data generated with the parameters learned from the previous experiments (Table 3) and real data collected from Microsoft academic search.

Method

To generate the citation graph which represents which authors cite which resources in their paper, we first parsed the publication information of all the AI researchers listed in Microsoft academic search[10]. This gave us a list of 266,101 authors, along with the number of publications each has produced. Then, we randomly sampled 400 authors[11] and generated the corresponding number of synthetic papers for each author. After that, we constructed the citation graph using the rich-get-richer model [13]. In this model, we first randomly sorted the papers; then, we created the citation for the papers sequentially. We assumed each paper cites 29 papers (the mean number of citations for 10 randomly sampled papers was 28.8). For each citation, we randomly cite one of the previously processed papers with probability p_{cite} and

[10]http://academic.research.microsoft.com/

[11]Our earlier experiments sent roughly this many emails.

randomly cite a paper cited by a previously processed paper with probability $(1 - p_{\text{cite}})$. We report experimental results with $p_{\text{cite}} = 0.5$.[12] This process ensures that the paper citations followed the power law. After the citation graph was generated, we randomly sampled 100 papers as the resources in the community. Based on data collected from our earlier experiments, we mark with 0.45 probability that the author of a citation really used that resource.

Based on the citation graph and the contribution probabilities we collected from our previous experiments (Table 3), we simulated the requests sent out by the community. We compare three methods for issuing requests:

1. **Random**: Send out requests that map the authors to the resources randomly.
2. **Greedy**: Based on the citation information, assign each author to the resource to which they are most likely to contribute.
3. **Decision-theoretic Optimization**: Issue requests using Algorithm 1.

We assumed the utility of the contributions of each task is $\log[100\mathbf{C}_t + 1]$, where $\mathbf{C}_t$ is the number of contributions that are made to task t. We chose this utility function because it has the property of diminishing utility for each additional contribution, a reasonable assumption since our community does not benefit from having all the contributions concentrated on only a few resources. We added 1 inside the log function to ensure nonnegative utilities. Since the expected utilities for many resources were less than one, we also multiply by 100 so that the utility is not dominated by the added constant factor.

We note that the Greedy and Random baselines are the strongest we could reasonably produce. For these baselines, we assign the authors with the highest probability of contributing to some resource first. Additionally, we break ties randomly, which has the effect of distributing contributions and dramatically improving the resulting utility.

Results

We generated 100 graphs using the method described in the previous section and simulated the requests sending in three different conditions: Random, Greedy, Decision-theoretic Optimization. The average expected utility of the three conditions on the five graphs are reported in Figure 6. The expected utility of the decision-theoretic algorithm is significantly higher than both baselines (using a two-tailed independent samples t-test). In particular, after issuing 400 requests, its expected utility is significantly higher than Random (58.9 v.s. 3.1, $p < 0.001$), it also performed significantly better than a strong baseline which assigned the authors greedily to the resources they were most likely to contribute to (58.9 v.s. 54.4, $p < 0.001$). Importantly, the figure also shows that decision-theoretic optimization needs to issue only 55% (220) of the 400 requests in order to reach the maximum expected utility of the best (greedy) baseline.

[12]We found that our results improve as we decrease p_{cite}, so we chose 0.5 as a representative value.

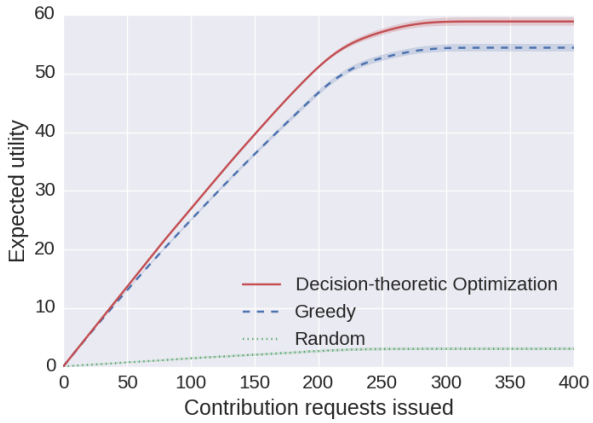

Figure 6. Decision-theoretic optimization achieves significantly higher expected utility and requires 55% as many requests (220) to match the maximum expected utility of the strongest baseline. Plot shows the mean expected utility over 100 simulations, with shaded 95% confidence intervals.

In addition to this main result, we performed a sensitivity analysis that showed our results to be robust and unchanged when we are not given the true probabilities of contribution. In this analysis, we provided the algorithms with access to the expected probability values used in our experiment, but sampled the true values from normal distributions centered at those values (and truncated at 0 and 1). Increasing the variance of these distributions until 2 standard deviations equaled the probability value itself did not alter our findings. Compared to the best (greedy) baseline, the decision-theoretic optimization method still achieved significantly higher expected utility after 400 requests (59.4 vs. 54.9, $p < 0.001$) and required only 56% as many requests to reach the maximum expected utility. This result is promising for a real-world deployment, where actual probability values would be drawn from a distribution rather than the expected value of that distribution.

LIMITATIONS & IMPLICATIONS

We are encouraged by the positive results from our experiments, but caution that there are several limitations to our study and our proposed method.

First, we only tested our method in one community. It may be the case that Open AIR is the best case for our datamining method, given the quality of corresponding data in Google Scholar. While more experiments are needed to demonstrate that our approach generalizes, there are other community where the approach has worked or is worth considering:

- **AirBnB / Craigslist:** AirBnB reputedly bootstrapped their inventory of rental properties by crawling Craigslist for candidate homeowners who had listed properties, sending them emails from supposed AirBnB "fans." [30]. While this is a blackhat example that may have violated Craigslist terms of service, it illustrates the method's applicability.

- **SummitPost / CascadeClimbers:**[13] The SummitPost community maintains an online guidebook of mountain-climbing route descriptions for a worldwide audience. CascadeClimbers is a regional community Website where climbers post pictures, accident updates and trip reports describing their outings. In contrast to the AirBnB / Craigslist example, these communities are complementary, not competitive. Since someone who has written a trip report has the knowledge to turn their tale into a more comprehensive and instructive route description, one might attract SummitPost contributions through the CascadeClimbers forum reply feature. For full coverage, one would wish to mine other regional sites as well.
- **500px / Reddit:**[14] 500px caters to a community of professional photographers who wish to showccase and sell their work to stock photo buyers. Reddit is a hierarchically-organized social news site; the photography subreddit features a wide ranging conversation about images and techniques. Someone who has posted several of their pictures on reddit is, therefore, a reasonable candidate member for 500px and might welcome a private message suggesting the site.
- **Movie Reviews / Twitter:** One might consider bootstrapping a movie review website by parsing Twitter posts for movie hastags and using text mining techniques to see whether the author expressed strong sentiment indicating proclivity to posting a review. While this example satisfes our requirement for "datamining bootstrappability," Twitter is noisy enough that such an approach seems unlikely to work.

The success of our recommended bootstrapping approach depends not just on the three conditions in our definition, but also on empirically determined parameters, such as the number of candidate nonmembers that can be unearthed via datamining, the accuracy of targeting and resulting response rate, the quality of the contributions, and the utility derived by the community over time. Further empirical studies are needed to determine how general is our method.

In addition, the performance guarantee of Algorithm 1 is based on the monotone submodularity of the utility function. Our algorithm might not perform as well when the community has a utility function with different properties. For example, a community might want contributions to focus on a few resources so resource popularity can attract newcomers to the community. However, the goal of our paper is not to provide a definite algorithm that can apply to every online community. Instead, we are trying to establish a decision-theoretic framework which allows the community designer to design their own algorithm that maximizes the community's utility. In the future, we plan to work with other online communities and come up with different optimization algorithms based on their individual utility functions.

[13]Respectively, at `www.summitpost.org` and `cascadeclimbers.com`.
[14]See `500px.com` and `www.reddit.com/r/photography`.

CONCLUSIONS AND FUTURE WORK

In this paper, we define bootstrapping an online community as a decision theoretic optimization problem. Although an optimal solution to the problem is combinatorially prohibitive, we present an efficient greedy algorithm and prove that it allocates requests within a constant factor of optimal. To demonstrate the practicality of our approach, we consider Open AIR, a newly created community for researching and reviewing open AI software and data resources. We show that text mining techniques, applied to Google Scholar pages, can extract several strong features that correlate with a person's interest in contributing a review.

Specifically, our results show that people who have authored a paper that cited an article describing a resource are more likely to comment on the resource than people who did not. Furthermore, by mining the context of these citations, our system can detect people who actually *used* the resource in their work; these people are significantly more likely to comment on the resource than people who simply acknowledge the resource as related work. Although we expected that strong positive or negative sentiment in the citation context would also signal a greater willingness to comment, the evidence did not support this conclusion.

Furthermore, our study shows that effective request design is an essential factor when encouraging non-members to contribute. Specifically, we found that first asking people a simple request (e.g., a binary "yes/no" question like "Would you recommend the resource?") significantly increased the likelihood that they would contribute the more time-consuming full-text review. This finding confirms the usefulness of the *foot-in-the-door* technique [16] in the context of bootstrapping an online community.

In the course of our experiments we were also able to learn parameter values for the conditional probabilities needed by our decision-theoretic model. Based on this information, we ran a simulation experiment showing that decision-theoretic control achieves comparable expected community utility while issuing only 55% as many contribution requests, compared to a strong baseline approach.

We are now ready to deploy the model on an even larger-scale study to see if we can complete the bootstrapping process and bring Open AIR across the tipping point to self-sustaining traffic. Additional important directions for future research include adding explicit budget constraints to our model, considering a wider range of utility functions, and applying our method to other online communities. Since the pairing of Google Scholar and Open AIR may represent the best case for datamining-based bootstrapping, further experimentation will help demonstrate the generality of our approach.

ACKNOWLEDGMENTS

We thank Chris Lin, Chien-Ju Ho, and the anonymous reviewers for especially helpful comments. This work was supported by NSF grant IIS-1420667, the WRF/Cable Professorship, a gift from Google, and the Allen Institute for Artificial Intelligence.

REFERENCES

1. Maneesh Agrawala and Chris Stolte. 2001. Rendering Effective Route Maps: Improving Usability Through Generalization. In *SIGGRAPH '01*.
2. Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. 2006. Group formation in large social networks: membership, growth, and evolution. In *KDD '06*.
3. Gerard Beenen, Kimberly Ling, Xiaoqing Wang, Klarissa Chang, Dan Frankowski, Paul Resnick, and Robert E Kraut. 2004. Using social psychology to motivate contributions to online communities. In *CSCW '04*.
4. Moira Burke, Robert Kraut, and Elisabeth Joyce. 2009a. Membership claims and requests: Conversation-level newcomer socialization strategies in online groups. *Small group research* (2009).
5. Moira Burke, Cameron Marlow, and Thomas Lento. 2009b. Feed Me: Motivating Newcomer Contribution in Social Network Sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. 10.
6. Brian S Butler. 1999. When is a group not a group: An empirical examination of metaphors for online social structure. *Social and Decision Sciences* (1999).
7. Robert B Cialdini and Noah J Goldstein. 2004. Social influence: Compliance and conformity. *Annual review of psychology* (2004).
8. Dan Cosley, Dan Frankowski, Loren Terveen, and John Riedl. 2006. Using intelligent task routing and contribution review to help communities build artifacts of lasting value. In *CHI '06*.
9. Dan Cosley, Dan Frankowski, Loren Terveen, and John Riedl. 2007. SuggestBot: using intelligent task routing to help people find work in wikipedia. In *IUI '07*.
10. Dan Cosley, Shyong K Lam, Istvan Albert, Joseph A Konstan, and John Riedl. 2003. Is seeing believing?: how recommender system interfaces affect users' opinions. In *CHI '03*.
11. Peng Dai, Mausam, and Daniel Weld. 2010. Decision-theoretic control of crowd-sourced workflows. In *AAAI '10*.
12. Chrysanthos Dellarocas and M Zhang. 2005. *Using online ratings as a proxy of word-of-mouth in motion picture revenue forecasting*. Technical Report.
13. David Easley and Jon Kleinberg. 2010. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA.
14. Andrea Esuli and Fabrizio Sebastiani. 2006. Sentiwordnet: A publicly available lexical resource for opinion mining. In *LREC '06*.
15. Marshall L Fisher, George L Nemhauser, and Laurence A Wolsey. 1978. An analysis of approximations for maximizing submodular set functions-II. In *Polyhedral combinatorics*. Springer, 73–87.
16. Jonathan L Freedman and Scott C Fraser. 1966. Compliance without pressure: the foot-in-the-door technique. *Journal of personality and social psychology* (1966).
17. Krzysztof Gajos and Daniel S Weld. 2004. SUPPLE: automatically generating user interfaces. In *IUI '04*.
18. Daniel Golovin and Andreas Krause. 2011. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research* 42, 1 (2011), 427–486.
19. Nicolas Guéguen. 2002. Foot-in-the-door technique and computer-mediated communication. *Computers in Human Behavior* (2002).
20. F Maxwell Harper, Dan Frankowski, Sara Drenner, Yuqing Ren, Sara Kiesler, Loren Terveen, Robert Kraut, and John Riedl. 2007. Talk amongst yourselves: inviting users to participate in online conversations. In *IUI '07*.
21. Eric Horvitz, Paul Koch, and Johnson Apacible. 2004. BusyBody: creating and fielding personalized models of the cost of interruption. In *CSCW '04*. ACM, 507–510.
22. Meiqun Hu, Ee-Peng Lim, Aixin Sun, Hady Wirawan Lauw, and Ba-Quy Vuong. 2007. Measuring article quality in wikipedia: models and evaluation. In *CIKM '07*. ACM, 243–252.
23. Ece Kamar, Ashish Kapoor, and Eric Horvitz. 2013. Lifelong Learning for Acquiring the Wisdom of the Crowd. In *IJCAI '13*.
24. Robert E Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren, and John Riedl. 2012. *Building successful online communities: Evidence-based social design*. MIT Press.
25. Claudia López and Peter Brusilovsky. 2011. Adapting Engagement E-mails to Users' Characteristics. CEUR Workshop Proceedings.
26. George A Miller. 1995. WordNet: a lexical database for English. *Commun. ACM* (1995).
27. George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. 1978. An analysis of approximations for maximizing submodular set functions–I. *Mathematical Programming* 14, 1 (1978), 265–294.
28. Paul J Resnick, Adrienne W Janney, Lorraine R Buis, and Caroline R Richardson. 2010. Adding an online community to an internet-mediated walking program. Part 2: strategies for encouraging community participation. *Journal of medical Internet research* (2010).

29. Jean-Charles Rochet and Jean Tirole. 2003. Platform competition in two-sided markets. *Journal of the European Economic Association* (2003), 990–1029.

30. Matt Rosoff. 2011. Airbnb Farmed Craigslist To Grow Its Listings, Says Competitor. *Business Insider* (May 2011).

31. S. Russell and P. Norvig. 2010. *Artificial Intelligence: A Modern Approach, 3rd ed.* Prentice Hall.

32. Jacob Solomon and Rick Wash. 2012. Bootstrapping wikis: developing critical mass in a fledgling community by seeding content. In *CSCW '12*.

33. Long Tran-Thanh, Sebastian Stein, Alex Rogers, and Nicholas R Jennings. 2014. Efficient crowdsourcing of unknown experts using bounded multi-armed bandits. *Artificial Intelligence* (2014).

APPENDIX: CONTRIBUTION REQUEST EMAIL

Subject: Did you use [[Resource Name]]?

Hi,

We'd love to hear your opinion on [[Resource Name]]! Click one of the links below to tell us your opinion about this resource.

I would recommend this resource to other AI researchers

I would NOT recommend this resource to other AI researchers

I haven't used [[Resource Name]]

Your opinion will be publicly posted, along with your name, on Open AIR, an open source collaboration hub for AI researchers run by the Allen Institute for Artificial Intelligence (AI2). Your contribution will help make Open AIR a valuable resource for the entire AI community.

Your response will also help us improve our understanding of how to encourage contribution to an online community. This study is being conducted by researchers at the [[anonymized author information]] who, in collaboration with AI2, are studying ways of bootstrapping online communities, including an analysis of user behavior in response to different email campaigns. Click here for more details about our study. This study is completely anonymous, but if for any reason you do not wish to participate, please click here and no data will be recorded. If you have any questions or suggestions, please email [[anonymized author information]]