# Sangeet Swara: A Community-Moderated Voice Forum in Rural India

**Aditya Vashistha**[†] **Edward Cutrell**[§] **Gaetano Borriello**[†] **William Thies**[§]

[†]University of Washington
{adityav, gaetano}@cs.washington.edu

[§]Microsoft Research India
{cutrell, thies}@microsoft.com

## ABSTRACT

Interactive voice forums have emerged as a promising platform for people in developing regions to record and share audio messages using low-end mobile phones. However, one of the barriers to the scalability of voice forums is the process of screening and categorizing content, often done by a dedicated team of moderators. We present Sangeet Swara, a voice forum for songs and cultural content that relies on the community of callers to curate high-quality posts that are prioritized for playback to others. An 11-week deployment of Sangeet Swara found broad and impassioned usage, especially among visually impaired users. We also conducted a follow-up experiment, called Talent Hunt, that sought to reduce reliance on toll-free telephone lines. Together, our deployments span about 53,000 calls from 13,000 callers, who submitted 6,000 posts and 150,000 judgments of other content. Using a mixed-methods analysis of call logs, audio content, comparison with outside judges, and 204 automated phone surveys, we evaluate the user experience, the strengths and weaknesses of community moderation, financial sustainability, and the implications for future systems.

### Author Keywords

HCI4D; ICT4D; Interactive Voice Response; IVR; India

### ACM Classification Keywords

H.5.2 Information Interfaces and Presentation (e.g., HCI)

## INTRODUCTION

Of the world's 7 billion mobile phone subscriptions, 78% are in developing countries [4], where phones are still used primarily for voice calls. Though smartphone penetration is growing – in India it is 10% [18], versus a mobile penetration of 73% [6] – there is still limited availability of Internet-enabled devices and bandwidth in rural contexts. Moreover, language and literacy constraints prevent usage of traditional online resources. In India, 26% of adults are illiterate [2] and 72% are illiterate with respect to English [9].

In light of these constraints, researchers have turned to *voice forums* as an inclusive means of accessing, reporting, and sharing information in rural communities. Using Interactive Voice Response (IVR), a voice forum allows users to record and listen to messages in their local language via ordinary calls on low-end mobile phones. Recent voice forums have spanned various domains, including citizen journalism [3, 13, 19], agriculture [1, 5], feedback on school meals [12], job search [30], and rural information portals [7]. Together, these forums have attracted millions of calls and hundreds of thousands of recordings.

However, one bottleneck that has prevented voice forums from rivaling the scale of large Internet websites is the process of content curation and moderation. In order to ensure respectful, accurate, and high-quality recordings, large voice forums typically employ a dedicated team of moderators, who screen recordings, offer feedback to contributors, and perform tagging, annotation, and categorization of posts. For example, both CGNet Swara [19] and Gram Vaani [3] currently employ 10-15 full-time moderators. If these platforms were to grow by orders of magnitude, it would be very difficult to manage commensurate growth of the moderator team.

In this paper, we describe Sangeet Swara[1]: a voice forum that uses *community moderation* to overcome the limitations of a dedicated moderator team. Sangeet Swara is a platform for people in rural India to share songs, poems, jokes, and other cultural content. In addition, it relies on users to rate the quality of content they hear on the system. These ratings, in turn, influence the order that recordings are played to other listeners, thereby improving the overall user experience. While community moderation has been successfully used on Internet websites, such as Reddit and StackOverflow, to date it has not been used to influence the playback priority on a voice forum. Extending community moderation to an IVR platform involves several unique challenges, including the limited affordances of the interface and users' limited experience with technology, especially in rural India.

Our primary contribution is the design and 11-week deployment of Sangeet Swara, which we evaluate along three dimensions: the engagement of users, the accuracy of community moderation, and financial sustainability. We find that users were highly engaged, with over 25,000 calls and 5,000 recordings from over 1,500 people. The service found unexpected uptake among visually impaired users, who were especially passionate about building and maintaining the community. We show that community moderation was 98% accurate in categorizing the content and gender of posts; it also made meaningful distinctions between high-quality and low-quality

[1]Sangeet Swara is a Hindi phrase meaning "musical voices."



*CHI 2015*, April 18–23 2015, Seoul, Republic of Korea

ACM 978-1-4503-3145-6/15/04 ... $15.00
http://dx.doi.org/10.1145/2702123.2702191

posts, and made judgments that were in 90% agreement with researchers on a sample of recordings. We also conducted an automated phone interview with 204 users, and offer qualitative findings regarding their perceptions of the service as well as the strengths and limitations of community moderation.

As a secondary contribution, we also advance the dialogue surrounding financial sustainability of voice forums in the developing world. Up until now, voice forums have relied on expensive toll-free lines in order to make them accessible to low-income callers. However, as usage scales, toll-free lines become too expensive to sustain [25]. The most direct solution to this problem is for users to pay for their own calls: an experiment that we tried in two different contexts. In Sangeet Swara, we eventually disabled the toll-free access, but found that even the most fervent rural users were unable to bear the cost of phone calls. As a follow-up experiment, we also describe Talent Hunt: an adaptation of Sangeet Swara to a higher-income context. Though Talent Hunt received considerable use, including about 27,000 phone calls from 12,000 different callers, this usage was driven mainly by a promotional contest and disappeared as soon as prizes were awarded. We discuss the lessons learned from this deployment and the implications for future voice forums.

## RELATED WORK

As discussed previously, several voice forums have targeted developing regions [1, 3, 5, 7, 12, 13, 19, 30]. Most closely related to our work is a system called Gurgaon Idol, where callers to an IVR service recorded and rated audio recordings to influence eventual playback on a community radio station [15]. However, this research focused on the usability of the recording interface, the voting interface, and training of users. It did not aim to incorporate the ratings to control the playback order in a voice forum.

Prior research has demonstrated that entertainment content drives technology adoption by low-income people in the developing world [24, 27]. In fact, even voice forums that are intended for other purposes often see many recordings of songs, religious verses, and other performances [5, 13, 19]. Recognizing the appeal of entertainment, a platform called Polly allowed people to record messages, manipulate recorded messages, and forward them to friends [26]. Polly spread rapidly to many users in Pakistan and has since provided job search information as well.

The curation tasks performed by callers on Sangeet Swara are related to crowdsourcing efforts in developing regions. Jana (formerly known as txtEagle [11]), mClerk [14], and MobileWorks [21] enable users to earn money by completing small tasks on low-cost mobile phones, using either mobile Internet or SMS. We are unaware of any paid crowdsourcing platform that administers tasks via an IVR interface, as used by callers to moderate posts on Sangeet Swara.

Several news and social networking sites, including Reddit, Slashdot, StackOverflow, and others, use community votes to determine the ranking of user-generated content. There is also a mature literature on collaborative filtering and community moderation algorithms [28]. However, we are unaware of prior efforts to perform community moderation on a voice forum, which is different in several ways. For example, audio content is more difficult to skim than textual content, meaning that users may lose patience in hearing and ranking lower-ranked posts. An IVR system can also track exactly what a user listened to and what content they skipped, which is difficult to do on a webpage. Finally, the limited affordances of an IVR interface and the limited technology skills in developing regions add more constraints to our design scenario.

## RESEARCH QUESTIONS

Our research goal is to create a vibrant virtual community that is inclusive of low-income users in rural India. Our success metrics are three-fold. First, do participants value their interactions with the community? Second, can the community moderate itself without outside assistance? And third, can it be financially sustainable?

In the following sections, we describe the design of Sangeet Swara, its deployment in rural India, and its evaluation using the three metrics above. We also present a follow-up experiment, Talent Hunt, that attempted to improve on financial sustainability by seeding usage in urban areas.

## SANGEET SWARA DESIGN

Sangeet Swara was a voice-based platform, accessible via mobile phone, where callers could record songs, poems, jokes, and other cultural content. A key aspect of the system is that it *ranked* the recordings based on feedback from the community. The ranking aimed to order the posts according to what was most likely to be enjoyed and appreciated by listeners. There was a single, global ranking computed across all recordings and all listeners in the system. In addition to the rank order, the system calculated a separate *playback order* that determined which post a listener heard at a given time. The playback order balanced the interests of listeners (who desired to hear high-quality posts) with the interests of content contributors (who desired to have as large of an audience as possible). Both the rank order and playback order were dynamically updated based on listeners' ratings of the content. We give more details on these orderings later.

### Call flow

Sangeet Swara relied on key press (DTMF) navigation: users listened to an audio menu and indicated their selection by pressing a digit. Though not as expressive as free-form speech, DTMF interactions have been shown to be robust and also preferred among users in rural India [23].

The high-level call flow is illustrated in Figure 1. The first thing that callers heard was an 8-second folk music excerpt, followed by a greeting in a male voice: "Friends, welcome to Sangeet Swara. You can record and listen to songs, poems, and jokes. Please note, it is free to call on this number." Then they were asked to select between the following options:

1. **Check on your posts.** Users who had recorded at least one post could listen to all of their recordings and also learn what rank they had obtained in the system. For users who had not recorded anything, this option was omitted.

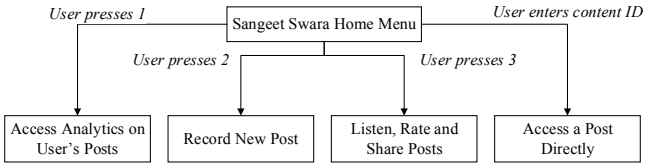


**Figure 1. High-level call flow of Sangeet Swara.**

2. **Record a post.** Users were encouraged to introduce themselves as part of their recording. We restricted the length of recordings to 60 seconds (plus a 10-second buffer for the introduction). After recording a post, users received an SMS with a unique five-digit numeric ID for that post. This ID could be shared and used to jump directly to the post (details below).

3. **Listen to other posts and rate them.** When users chose to listen to posts, first we played the top ranked post (introduced as "the best message on the basis of community votes"). Playing the best message first ensured that callers heard at least one high-quality recording per call. It also engendered friendly competition to be featured in the top spot. After the first post, users listened to other posts in the playback order computed by the system. In advance of playing each post, the system announced its current rank among all posts recorded to date. Since the rank order and playback order were different, users listened to an unpredictable mixture of highly-ranked and low-ranked content.

   After listening to a recording, users were required to give feedback by pressing a key for "like" or "dislike." Users could also interrupt the playback of a post to offer an early judgement, in which case the remainder of that post was skipped. Each user had only one vote to count towards a given post; if they played a post twice, they could change their vote but not increase it. Users could also press a key to receive an SMS that was suitable for sharing with friends. The SMS contained the unique ID of the post and instructions for accessing it on Sangeet Swara.

4. **Jump directly to a post.** Users could directly jump to a post by entering its ID number at the main menu. The first digit of the ID number was always different than the other options at the main menu, enabling users to make the jump immediately without navigating through any other menus.

Our design of menu prompts respected the lessons learned from prior IVR systems in low-income communities [5, 8, 16, 17]. The prompts were recorded in the local language and accent of the target area (North Indian Hindi), with slow and clear diction by the speaker (a male). Prompts explained each possible action before the corresponding key press; keys had consistent meanings across all menus; multi-digit inputs were avoided as much as possible; and invalid key presses led to explanatory error messages.

We used iterative prototyping to refine the system in advance of deployment. In a formative evaluation with our colleagues, 28 callers placed 236 calls over a period of 3 weeks. To understand usability barriers, we performed participant observation and conducted five unstructured interviews. This led to several improvements. For example, before posting a recorded message, the system played it back and asked for confirmation from the user.

Towards the end of our field deployment, we also augmented the call flow with an additional feature. We identified regular users (those who had called at least ten times) and notified them, at the beginning of the call, that they were now a "senior member" of the Sangeet Swara community. Commensurate with this distinction, we asked them to take on a new responsibility, which was to answer one pre-recorded question at the beginning of each phone call. As detailed later, we used these questions both to conduct surveys of the users and to take users' help in categorizing the content recorded by others. The survey questions solicited free-form audio responses, while the categorization questions were multiple choice. Senior members could advance to the main menu of Sangeet Swara only after answering the question posed.

#### Rank Order

The rank order aims to sort posts by increasing order of quality, as determined by users' upvotes and downvotes. There are two criteria that contribute to the rank ordering:

- High scores: a post with a higher ratio of upvotes to downvotes is likely to be of higher quality.
- High confidence: for comparable ratios of upvotes to downvotes, we have more confidence that a post is good if more people have voted on it.

Following Reddit's algorithm for sorting comments [20], we integrate both of these concerns by calculating the lower bound of the Wilson score confidence interval for a Bernoulli parameter:

$$\frac{\hat{u} + z^2_{\alpha/2}/2n - z_{\alpha/2} * \sqrt{[\hat{u}\,(1-\hat{u}) + z^2_{\alpha/2}/4n]/n}}{1 + z^2_{\alpha/2}/n}$$

Here, $\hat{u}$ is the fraction of upvotes, $n$ is the total number of votes, and $z_{\alpha/2}$ is the $(1-\alpha/2)$ quantile of a standard normal distribution. We used the lower bound of a 95% confidence interval ($\alpha = 0.05$) to compute the rank score. The post with highest score was assigned the top rank.

#### Playback Order

The playback order refers to the sequence in which a user listens to posts. The playback order needs to balance the following competing criteria:

- Listeners want to hear good content. This prioritizes posts with a large fraction of positive votes.
- Contributors want their posts to receive a fair ranking. This prioritizes posts with a small number of total votes (since more votes lead to a more accurate assessment of quality).

To balance these concerns, we calculate a post's playback priority according to the following formula:

$$\frac{U}{U+D} * (1 - DF)^{U+D}$$

Here, $U$ is the number of upvotes for the post, $D$ is the number of downvotes for the post, and $DF$ is a discount factor

that serves to balance the concerns of listeners and contributors. We used a discount factor of 0.333 after analyzing a range of values and their impact on example scenarios. We initialize posts with a single upvote to avoid division by zero.

The first term of the equation represents the priority of playback according to listeners (criterion 1), while the second term captures the priority for contributors (criterion 2). Larger numbers represent a higher priority.

When a user elects to listen to posts, the system plays the highest priority posts that the user has not yet voted on. If the user has voted on all the posts, then attention is restricted to posts the user has liked in the past, and playback proceeds in rank order instead of playback order.

Our calculation of priority is similar to other rankings that reconcile the competing metrics of quality and recency. In our context, recency corresponds to the total number of votes that a post has received to date, rather than the time elapsed since the recording.

## SANGEET SWARA DEPLOYMENT

Sangeet Swara was deployed for eleven weeks. In order to lower the barriers to participation, we launched the service using a toll-free (1-800) number. However, as toll-free lines become expensive at a large scale, we also wanted to explore if users would pay for the phone calls. Thus, we moved the service to a regular number after about seven weeks.

To create awareness about Sangeet Swara in rural and small-town India, we posted a message on CGNet Swara, a voice forum for citizen journalism that has considerable reach in rural areas. The post was accessible to CGNet Swara callers for two days, during which time it was heard by 393 unique callers. Out of those, 73 people placed a call to Sangeet Swara. In order to help Sangeet Swara feel familiar to prior users of CGNet Swara, and to set the standards for the community, we seeded Sangeet Swara with fifteen songs and poems that appeared previously on CGNet Swara.

Sangeet Swara had significant uptake. As summarized in Table 1, the system received over 25,000 phone calls from over 1,500 people. There were about 5,000 posts recorded, about 200,000 playback events, and about 140,000 votes cast. Figure 2 depicts the usage over time. As detailed in the next section, usage was highest among low-income people from rural and peri-urban areas of northern and central India. Also, the service saw a high uptake among visually impaired users.

Unfortunately, usage of the system dropped dramatically when we converted the toll-free lines to regular lines. We revisit the question of financial sustainability later, both in the context of Sangeet Swara as well as a follow-up experiment.

## EVALUATING THE USER EXPERIENCE

This section aims to evaluate users' experience of Sangeet Swara, including the worth they attached to the system.

### Methods

We used a mixed methods approach spanning qualitative and quantitative analyses. Our primary tool was an automated phone survey, which presented a single pre-recorded question to regular users each time they called. There were nine questions about basic demographic data (sex, age, education, technology exposure, etc.) and six open-ended questions probing the background of the listener (e.g., "tell us about yourself"), their conception of Sangeet Swara (e.g., "how would you describe Sangeet Swara to a friend"), the quality of community moderation, and the strengths and weaknesses of the platform. The survey was live for ten days. All responses were provided in audio format and then translated, transcribed, and analyzed using open coding and axial coding. A total of 204 people (out of 409 regular users) answered one or more of the survey questions, and each question was answered by at least 100 people. On average, free responses were 36 words long.

We performed a content analysis of 100 randomly selected posts. When we learned of the prevalence of visually impaired users, we conducted ten semi-structured telephone interviews to understand their experience in more detail. We used open coding to analyze the data obtained from the content analysis, the live interviews, and the automated interviews. We also studied call logs to understand usage patterns.

| | Sangeet Swara | Talent Hunt |
|---|---|---|
| Deployment Duration | 11 weeks | 22 weeks |
| Language of Prompts | Hindi | English |
| Calls | 25,381 | 27,514 |
| Callers | 1,521 | 11,751 |
| Posts | 5,376 | 368 |
| Contributors | 516 | 304 |
| Average Call Duration | 4.96 mins | 2.37 mins |
| Total Plays of Posts | 198,898 | 42,383 |
| Upvotes | 40,590 | 10,832 |
| Downvotes | 99,150 | 3,375 |
| Share Requests | 773 | 251 |
| Direct Jumps to Post ID | 7,871 | 11,868 |

**Table 1. Usage statistics for Sangeet Swara and Talent Hunt.**

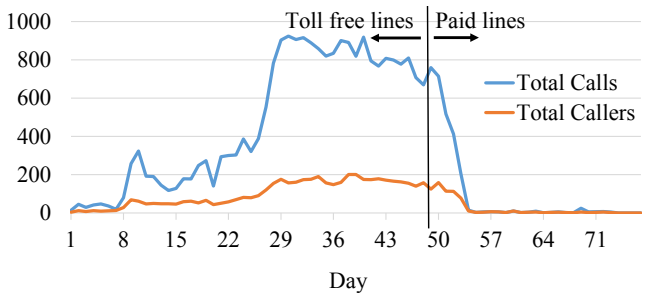


**Figure 2. Call statistics for Sangeet Swara.**

### Results

Sangeet Swara led to high levels of community engagement, with users becoming devoted champions of the system. Similar to other platforms for user-generated content [22], the top 10% of callers placed 70% of total calls. The top 10% of content authors were responsible for 60% of all messages.

### *User Analysis*[1]

Surprisingly to us, Sangeet Swara found broad and impassioned usage by visually impaired users; of those users who told us "something about themselves," 26%[2] voluntarily disclosed that they were visually impaired. The uptake by visually impaired users was organic; though voice is a natural mode of interaction for the visually impaired, we did not anticipate this usage and played no role in promoting it. Our interviews with them led to a broader study on their creation, consumption, and sharing of educational content [29].

Our users were predominately young men: 94% were male (average age=25 years) and 6% were female (average age=22 years). The youngest user was an eight year old and the oldest user was a 52 year old man. Users came from a broad range of educational backgrounds: 16% held or were pursuing a Master's degree, 40% held or were pursuing a Bachelor's degree, 24% were in high school, and 17% were in middle school. Two respondents were in primary school, and one described himself as uneducated. Our users came from a variety of vocations: 54% were students, 17% were teachers, 6% were working in private jobs, 5% were unemployed, 4% were musicians and 4% were farmers. The 42% of users who were employed reported a median annual income of USD 960 with a maximum of USD 7000. About half of users were from rural areas or small towns, while others were from larger cities.

SMS use was fairly common, with 61% reporting its use. However, most respondents had little experience with the Internet: only 16% had used an email account at least once, and the same fraction had used Facebook. Many users had never heard of Facebook; three people said: *"We don't have a Facebook account but we have an account in Bank of India."*

### *Content Analysis*

Our open categorization of 100 randomly sampled posts found generic messages (N=36), songs (N=21), poems (N=16), users' introduction about themselves (N=6), songs played from another playback device (N=6), instrumental performances (N=4), jokes (N=3), blank messages (N=3), questions (N=3) and current news (N=2). Eighty-eight messages were recorded by males, two were recorded by females and one message was recorded by a group consisting of both males and females. We did not categorize the gender of blank messages and messages that recorded playback from other devices. Fifty-six people reported their location while recording the messages. Most of the messages came from Madhya Pradesh (N=25), Rajasthan (N=11) and Uttar Pradesh (N=8).

Songs and poems accounted for about half of the content, and were in a variety of styles. The songs spanned recognizable hits, folk music, and original pieces; solos and duets; a cappella and pieces with instruments. The top 50 posts (analyzed in a later section) are available for listening at `http://soundcloud.com/sangeet-swara/`.

The other half of posts emerged as general social media. Many messages were wishing well to the 'friends' users made on Sangeet Swara. There were greetings, good morning messages, and good night messages for other users, and responses from one user to another. Many messages were about recent topics of national or regional interest. For example, there were five messages about the 2013 North India floods. Eight users gave their phone number and encouraged others to contact them for chatting. Male users often recorded compliments for female contributors, praising their beautiful singing and sometimes requesting their phone number.

While the vast majority of posts on Sangeet Swara were respectful in tone, we found and deleted 22 posts containing abusive language or derogatory comments. As seen in prior forums such as Avaaj Otalo [5], users took an active role in policing the system, e.g., by urging others to record cultural content and to avoid abusive comments. Though we did not have time to evaluate community-based flagging and deletion of unwanted posts, this feature would be important at scale.

### *Value Offered to Users*

Many users attributed great value to their interactions on Sangeet Swara. They recorded strong positive sentiments about the service and shared interesting anecdotes about how Sangeet Swara was impacting their lives. They considered it to be a platform where people show their creativity, voice their opinions, and record interesting content. This sentiment was often strongest among visually impaired users:

> *My mother and father are laborers. You are like my father, my god. I want to thank you again and again, this small kid wants to respect you from the bottom of my heart. I listen to abundant good content on Sangeet Swara. I never got the opportunity to hear such content elsewhere. I am in love with Sangeet Swara since the first day.*
>
> P1 (Male, Student, Visually impaired, 18 years, Madhya Pradesh)

Visually impaired people used Sangeet Swara to showcase their talent, build social capital, and share information. Some people considered Sangeet Swara to be a platform *"to learn and understand the principles of life."* Some people considered it to be a conduit for national peace and integrity, and believed that *"it is proving the mantra of India: Unity in diversity."* One user said:

> *I am blind so I couldn't get educated. I want to thank you from the bottom of my heart because you enabled all blind people to get in touch with each other and show our talent. No matter how much I praise, it won't be enough.*
>
> P2 (Male, Uneducated, Visually impaired, Madhya Pradesh)

Many participants thought of Sangeet Swara as a platform for promoting poor musicians from rural parts of India. They thought of it as a stage for such musicians to showcase and improve their talent, to overcome stage fright, and to step toward India-wide recognition and fame.

> *You can put your hidden talent on the forefront. People don't feel that anyone is listening and thus, they can perform without any hesitation. A performer will feel as if*

[1]Our survey targeted users who had called at least 10 times.

[2]Because different questions were answered by different numbers of users, we report the percentage of users answering a given question.

*he is alone but a lot of people listen to it later on Sangeet Swara. People with stage fright can present their talent on this wonderful platform. No one can mock you. You will get to meet new people. People in far-off locations will hear you.*

P3 (Male, Teacher, 26 years, Madhya Pradesh)

Many participants appreciated the voice medium, stating that it makes the process of information curation and dissemination much simpler than text-based alternatives. They considered Sangeet Swara to be an inclusive portal for low-literate people, visually impaired people, and tribal people in India.

*Sangeet Swara is trying to get talent from people in villages and towns. It is a channel for talented people who never got an opportunity to show their talent. Sangeet Swara is trying to get recognition and provide a channel for such people.*

P4 (Male, Musician, 22 years, New Delhi)

Sangeet Swara also helped people build self confidence. Three people reported feeling important when they use Sangeet Swara. Some people felt that Sangeet Swara was also playing a role in improving their grammar, vocabulary and communication skills.

*I get a lot of knowledge by using Sangeet Swara. Some people record questions, which increases our knowledge. We get to listen to things we have never heard. We learn new vocabulary and sometimes new accents as well. I feel great when people vote for me and give me feedback, be it a good feedback or bad. I consciously think of ways to improve my messages.*

P5 (Male, Student/Farmer, Uttar Pradesh)

## ANALYSIS OF COMMUNITY MODERATION

Our goal in this section is to assess the usability and effectiveness of community moderation in Sangeet Swara. We start with three quantitative approaches: evaluating the accuracy of crowd categorization tasks, comparing the top 50 posts to the bottom 50 posts, and comparing the ranking of posts to an "expert" ranking. Then we consider qualitative feedback from the users.

### Categorizing the Posts

Before making a judgement regarding the quality of a post, a basic task one might expect from a moderator is to categorize the content along various dimensions, such as the type of recording, gender of the contributor, language of the post, etc. Allowing listeners to search or filter content according to these metadata would be an important feature of a scalable voice forum, even though we did not implement such functionality in Sangeet Swara. Especially for Indic languages such as Hindi, it is very difficult for current speech recognition technologies to automate or assist with such tagging and categorization tasks.

As described previously, we designated regular users of Sangeet Swara as "senior members" and sometimes asked them to help categorize messages at the beginning of the phone call. Our content analysis revealed that becoming a "senior member" had a strong positive effect on users. The designation made them feel privileged, honored, and grateful. They felt more accountable for improving content quality, casting votes diligently and performing tasks:

| Task Type | Offered | Done | Response Rate | Accuracy |
|---|---|---|---|---|
| Content type | 1704 | 1551 | 91.0% | 98% |
| Gender | 2000 | 1895 | 94.7% | 98% |

**Table 2. Results of categorization tasks done by community.**

*I am now a special person on Sangeet Swara. I have to categorize posts. Please don't use any abusive language on this forum. Please don't say anything wrong because they have made me a senior member and if you do anything wrong then I will tell them.*

(Post on Sangeet Swara)

We asked users to categorize posts along two dimensions: content type and gender. To classify content type, users pressed a key to indicate if the recording was (1) a song, (2) a joke, (3) a poem, or (4) none of the above. To classify gender, users indicated if the speaker was (1) a male voice, (2) a female voice, or (3) they couldn't tell if it was male or female.

Senior members were asked to categorize the top 50 recordings. Whenever senior members called, one of the tasks was randomly presented to them. In total, 3704 categorization tasks were completed; for each post, we received at least 33 judgements of content type and 40 judgements of gender. The tasks were offered to 291 users, out of which 146 completed the task. The top 20% of workers performed 66% of the tasks.

For each categorization task, we aggregated the responses from the crowd and selected the majority answer as the community response. Before inspecting the community responses, a researcher categorized all the posts by the same criteria. We calculated the crowd's *accuracy* as the fraction of judgements that agreed with the researcher's and the *response rate* as the percentage of users who completed a task when it was offered to them.

The results appear in Table 2. The community showed high accuracy (98% agreement with researcher) on both content and gender classification. For each task type, only one message led to disagreement (a Bollywood hip-hop song for content type, and a muffled voice for gender). The response rate was 95% for gender, and 91% for content type. We conjecture that the gender task was easier, leading to more responses.

Many users recorded messages to share their feedback about the tasks. Some users requested more variety of tasks, while others recorded messages critiquing the recordings they categorized. Though the majority of users were excited about helping Sangeet Swara by performing tasks, two people recorded complaints. For example:

*I don't want to do any tasks. I just want to listen to the content right away.*

(Post on Sangeet Swara)

| | Content Type | | | | Gender | | | Inaudible | | Language | | Duration (seconds) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Song** | **Joke** | **Poem** | **Misc** | **Male** | **Female** | **Not Sure** | **Yes** | **No** | **Hindi** | **English** | **Mean** | **Median** | **Mode** |
| **Top 50** | 16 | 7 | 23 | 4 | 30 | 20 | 0 | 0 | 50 | 49 | 1 | 48 | 49 | 70 |
| **Bottom 50** | 10 | 0 | 2 | 38 | 46 | 0 | 4 | 4 | 46 | 48 | 2 | 40 | 35 | 70 |

**Table 3. Analysis of the top 50 and bottom 50 posts.**

### Top 50 vs. Bottom 50 Analysis

In order to probe whether the community applied consistent criteria for desired content, we compared the top 50 messages with the bottom 50 messages (out of a total of about 5,000 messages). If community moderation was successful, the desired content would rise to the top and the poor content would sink to the bottom. We analyzed the posts on several dimensions, including content type, gender, audibility, language, content duration and the geographic region of the caller.

The results of the comparison appear in Table 3 (with the exception of geography, which is presented later). We find a significant difference in content type in the top 50 and bottom 50 posts ($\chi^2(3, N = 100) = 53.5$, $p < 0.0001$). In the top 50 posts, only four posts were in the miscellaneous category as opposed to 38 such posts in the bottom 50. Most of the posts in the bottom 50 were personal messages for another user (N=15), information about other IVR services (N=7), comments on others' posts (N=5), and blank or nonsensical messages (N=4). This demonstrates that the community was successful in promoting songs, poems and jokes to top positions, while pushing messages deviating from the intended usage of Sangeet Swara to bottom positions.

We also found a significant difference in gender between the top 50 and bottom 50 posts ($\chi^2(2, N = 100) = 27.3$, $p < 0.0001$). Of the top 50 posts, 40% were recorded by females: twenty times the fraction of female recordings in our random sample of content (2%). In contrast, the bottom 50 posts did not contain any recordings by females. This trend corroborates our user and content analyses, in which most users were male and offered special attention, flirting, and adulation to female contributors.

The top 50 and bottom 50 posts did not show significant variations in language, duration, or inaudible posts. However, it is worth noting that the bottom 50 messages contained four inaudible posts while all of the top 50 posts were audible.

We also tabulated the approximate geographical location of callers based on their caller ID[3]. The majority of content authors belonged to similar locations in the top 50 and bottom 50 posts: Rajasthan (N=10, M=14), Madhya Pradesh (N=9, M=14), Uttar Pradesh (N=10, M=7) and Delhi (N=7, M=1).

### Community Ranking vs. Researcher Ranking

As an additional validation that community moderation resulted in a meaningful ranking of content, we compared the ranking of messages on Sangeet Swara to a ranking determined by a group of researchers. If these rankings differ, it does not prove that Sangeet Swara rankings are invalid, as the differences could be due to varying tastes of the demographic groups. However, if the rankings agree, it provides additional evidence that the community can perform its own moderation tasks without relying on outside assistance.

In order to compare the judgements of users and researchers, we restricted our attention to songs (the most frequent content type). Restricting attention in this way allowed a more direct comparison of quality, without conflating user preferences for one content type over another. Our experimental design asked researchers to compare a pair of songs, and to see if their preference matched the relative rank of those songs on Sangeet Swara. We prepared 20 pairs of songs from Sangeet Swara. The first ten pairs consisted of one song ranked in the top 20, and one song ranked in the middle 10. The second ten pairs consisted of one song ranked in the top 20, and one song ranked in the bottom 10. We randomized the order of the pairs, and the order of songs within pairs.

For each pair of songs, we asked three researchers (1 male, 2 female, Indian natives, average age=28 years) to select the one they liked more. Researchers were instructed to focus on the quality of the singing, doing their best to ignore any variations in language or (if the song is well known) any preference for the original version. Researchers were blind to the ranking of songs on Sangeet Swara, and rankings by each other. We used a majority vote to determine the researchers' ranking of a given pair. We compared the researchers' vote with the community ranking to measure agreement between them.

When comparing top-ranked and bottom-ranked posts, 90% of song pairs received the same ranking by researchers and Sangeet Swara users. This amount of agreement is unlikely to happen by chance (a binomial test of 10 trials, each with 50/50 chance of agreement, leading to at least 90% of judgements in either direction, yields $p = 0.02$).

When comparing the top-ranked posts and posts with middle ranking, only 60% of pairs received the same ordering from researchers and Sangeet Swara users. There are several possible interpretations of this result. The top and middle posts were more similar in quality, requiring more subtle distinctions. As song preferences are highly variable, we may have obtained a higher match if we had used a larger group of researchers. While we cannot rule out the possibility that Sangeet Swara users were less careful or less capable to compare the songs, this assertion is not supported by our other observations (such as the high accuracy on categorization tasks).

### Qualitative Views of Community Moderation

To understand users' feelings about community moderation, we included a question on this topic in our automated telephone survey. The question asked, *"When you listen to posts, Sangeet Swara tells you the rank of the post. Do you feel that good posts are ranked higher on Sangeet Swara and bad posts are ranked lower?"* We received 126 responses, which were transcribed, translated, and analyzed in two different ways.

[3]In India, a phone number reveals the telecom circle in which a SIM card was purchased, but not the circle where it is currently located.

The first analysis was a coarse-grained sentiment analysis. The largest category of responses (36%) were neutral or difficult to classify. However, 35% of respondents generally agreed that good content was ranked higher and bad content was ranked lower. A slightly lesser fraction (29%) were not satisfied with the quality of community moderation.

To understand users' views with more nuance, we analyzed the transcripts using open coding and axial coding, arriving at several themes. We found that many people understood that their votes decided the rank and also influenced the playback order. These users emphasized the need to vote honestly:

> *Some messages are really good and their rank is also good. However, around 10% messages aren't good and yet they have a good rank. It is not the fault of the system. The voters should understand which message should be taken to a high level and which to a low level. I would like to tell all listeners that they should listen to messages carefully and then vote honestly. Each vote is precious.*
>
> P6 (Male, Student, 19 years, Uttar Pradesh)

Many people agreed that the quality of community moderation is good and the rank of good quality content is generally higher than the rank of bad quality content. For example:

> *The good songs are higher ranked and the bad songs are lower ranked. I am happy that you decided to rank the posts by our votes.*
>
> P7 (Male, Student, 15 years, Jharkhand)

We found some people who believed that Sangeet Swara administrators decided the rankings. They did not understand how their votes influenced the rank and playback order:

> *Whatever rank the system chooses is right. The good messages have higher rank and the bad messages have lower rank. I trust you that you will never favor anyone. You will categorize the messages properly, give good rank to good messages and will depress bad messages.*
>
> P8 (Male, Telephone operator, Visually impaired, 42 years, Madhya Pradesh)

Some people didn't agree with the ranks assigned to posts, and were unhappy with the quality of community moderation. A few people put the blame on Sangeet Swara administrators for inappropriately assigning the rank:

> *I think you don't listen to the messages. Some messages are very good but have low rank and some messages are useless but they have good rank. Either you are confused or there is some fault somewhere in your system.*
>
> P9 (Male, Student, 19 years, Uttar Pradesh)

Others felt that careless voting by users is responsible for poor quality moderation:

> *Not all the messages appear to have the right rank. I think the reason behind that is voting by the community. I think at many places people do not vote responsibly. They just want to go ahead in the playback list and they don't care whether they are pressing 1 or 2.*
>
> P10 (Male, Government servant, Uttar Pradesh)

Eight respondents demonstrated a lack of understanding between "rank" and "playback order." For example:

> *The rank which is told at Sangeet Swara, I don't understand it. Sometimes the rank is 2511, and the next message is ranked 3303 and then the next to it is 1127. Sometimes it increases and decreases and I don't understand it.*
>
> P11 (Male, Teacher, Jharkhand)

Though the distinction between rank and playback order is necessary to ensure fair playback and voting policies, there could be better and simpler ways to communicate the rankings to users. For example, instead of reading the numeric rank, a prompt could say "this is a new post, and we really need your opinion", "this post is an old favorite", "this post is liked by some people, but more input is needed", and so on.

To summarize, though community moderation demonstrated effectiveness from a quantitative standpoint, our qualitative analysis reveals that there is room to improve on how the system is understood and appreciated by users.

## EVALUATION OF FINANCIAL SUSTAINABILITY

To create a voice forum that can scale and sustain without outside assistance, moderating the content is only part of the equation. The other challenge is financial sustainability. In particular, there needs to be a way to support the cost of phone calls as more and more users are calling the system.

Given how deeply many users seemed to value Sangeet Swara, we thought that a subset of users may be willing to pay for their own phone calls, thereby sustaining the system without external funding. Thus, after spending about 3,000 USD on seven weeks of toll-free support, we planned a switch to a regular line (which costed users the same as a normal phone call – 1 to 2 cents per minute, depending on their mobile plan). This required users to call a different phone number, which we announced as part of the welcome message for the five days preceding the change. We are confident that users understood the change in number, because the forum was immediately inundated with emotional requests to continue the toll-free service. For example:

> *I am very sad. Please, don't change the number. I fold my hands and request. Please consider my request. Not only me, everyone wants it to be toll-free. If it is a paid number then people won't be able to use it. Please don't reject our plea. Sorry sir. I fold my hands and pray, please don't change the number. Please cancel the announcement.*
>
> P12 (Male, Student, Visually impaired, 19 years, Uttrakhand)

Unfortunately, this student's prediction was correct. As illustrated in Figure 2, usage steeply declined without the toll-free lines. Within four weeks, it died out completely. We will say more about this result as part of our closing discussion.

## FOLLOW-UP EXPERIMENT: TALENT HUNT

Though users of Sangeet Swara derived significant value and meaning from the system, they were unable to pay for their own phone calls, which limited the scale we could achieve. As a follow-up experiment, we wanted to see if similar value could be delivered to a slightly higher-income group that might be able to afford to make phone calls without toll-free lines. If successful, such an experiment could grow into a very large ecosystem, giving more opportunities for monetization and cross-subsidization of low-income users.

To explore this idea, we deployed Talent Hunt: a forum of songs, poetry, jokes, and other cultural content targeting college students in urban India. The infrastructure and call flow for Talent Hunt was almost the same as Sangeet Swara, though prompts were recorded in English (arguably the most common language for college students in India) instead of Hindi. Moreover, the biggest difference was in the financial incentives offered. Instead of using toll-free lines, we promoted participation by awarding a smartphone (Nokia Lumia 710) to authors of top-ranked posts. We made one award per week for the first six weeks. After six weeks, we still announced one winner per week (and featured winners in the main menu), but did not award any phones. Our hope was that the material prizes would be sufficient to seed interest in the forum; future participation would be sustained by social recognition and attachment to the community. We promoted Talent Hunt using posters (in college campuses), email, social media, and targeted outreach by student volunteers.

The usage of Talent Hunt over time is illustrated in Figure 3. Though it received calls from 11,751 people (7 times more than Sangeet Swara), unfortunately this usage was driven entirely by the material awards. As soon as the last phone was awarded, participation dropped to zero. We ended up spending more on the phones (USD 1517) than the users collectively spent on the airtime (USD 1305, assuming $0.02 / min).

The award-based incentive structure also had deleterious effects on the quality of the voice forum. Despite the large number of callers, only 368 posts were recorded, which is 15 times less than on Sangeet Swara. Rather than building a supportive community of participation and sharing, users of Talent Hunt were often calling only to vote for a friend. Among ten of the top-rated posts, 99.5% of votes obtained were direct votes, in which the voter jumped directly to the post by entering its ID from the main menu: evidence that the caller was trying to support someone they knew in real life.

We also conducted semi-structured qualitative interviews with ten users who were among the top vote recipients. These users indicated that they strategically mobilized large groups of people to vote on their posts. Some of them made announcements in classrooms, on social media sites, and even at a wedding in order to gather more votes.

The quality and quantity of user engagement was also much lower than Sangeet Swara. By the standards established in Sangeet Swara, no member of Talent Hunt would have been designated a senior member. As most of the users were college students in tier 1 and tier 2 cities, they also had access to social media platforms such as Facebook, WhatsApp, Twitter, etc. As a result, Talent Hunt offered less value to them.

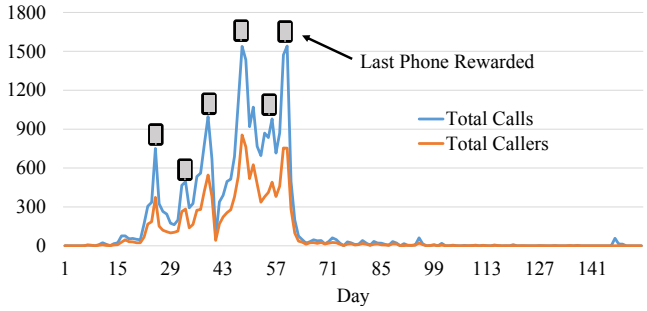


**Figure 3. Call statistics for Talent Hunt.**

To summarize, our experience with Talent Hunt showed that offering financial awards for top-ranked recordings brought several hazards to a community-moderated voice forum. Patterns of participation and voting became grossly distorted by users who sought to help their friends. Moreover, the incentive failed to seed long-term participation. Though users were able to pay for their own calls during the contest period, the usage expired as soon as the contest was over.

## DISCUSSION AND CONCLUSIONS

This paper describes lessons learned from Sangeet Swara, a community-moderated voice forum in rural India. The system evoked a passionate response from users, particularly those with visual impairment, who discovered and appropriated the platform without any outreach on our part. We believe that the ability to be an equal participant (and moderator) of a voice forum was a uniquely empowering experience for these users. Our study shows that this community of untrained callers can accurately perform their own moderation tasks, including categorizing and rating posts, thereby mitigating the bottleneck of a dedicated moderation team.

For community-moderated voice forums to scale further, they also require financial sustainability, which was not achieved by either Sangeet Swara or Talent Hunt. Given that Sangeet Swara users were passionate about using the system, their reluctance to pay for the phone calls is almost akin to an "impossibility proof": for users in this demographic, it is very difficult for a voice forum to deliver sufficient benefits for users to consider paying for the calls themselves. Conversely, in the case of Talent Hunt, we believe that users were able to pay, but had lesser interest in the service, since they had access to online social media platforms.

To reduce costs in the future, one promising approach is to transfer audio content via mobile data connections (as they become available) instead of voice calls [10]. It may also be possible for callers to perform more general audio microtasks, similar to Mechanical Turk but administered over IVR. The revenue generated could help offset the costs of calls.

Several aspects of community moderation remain untested. Sangeet Swara focused on the domain of entertainment, where the content is relatively uncontroversial. Extending to domains such as politics and citizen journalism will require

sensitivity to stronger disagreements between callers, which could impact their ratings as well as their flagging of posts for deletion. Also, 94% of Sangeet Swara users were male. It will be important to understand how to build a community that is inclusive and inviting towards women.

There are rich opportunities to broaden the scope of voice forums. Users of Sangeet Swara desired additional interactions, such as sending personal messages and listening to all posts by a given person. Generalizing a voice forum in this way could lead to a flexible social networking platform over IVR, leading to even greater uptake and engagement by rural users.

## ACKNOWLEDGMENTS

We are grateful to Kalika Bali, Sankalp Chhabra, Shubhranshu Choudhary, Deepti Desai, Nakull Gupta, Dhruv Parthasarathy, Indrani Medhi Thies, Satish Sangameswaran, and Sridhar Vedantham for all of their help. We also thank the anonymous reviewers for their helpful suggestions. This work was funded in part by the Noe Professorship at the University of Washington, as well as the Catalyst Center for Sustainable Development at Microsoft Research India.

## REFERENCES

1. Awaaz.de. **http://awaaz.de/**.
2. CensusInfo India 2011. **http://www.devinfolive.info/censusinfodashboard/**.
3. Gram Vaani. **http://gramvaani.org/**.
4. The World in 2014: ICT Facts and Figures. Tech. rep. **http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf**.
5. Avaaj Otalo: a field study of an interactive voice forum for small farmers in rural India. In *CHI* (2010).
6. Press Release on Telecom Subscription Data as of 31st March, 2014. Tech. rep., TRAI, 2014. **http://www.trai.gov.in/Content/PressDetails/2148_0.aspx**.
7. Agarwal, S. K., Kumar, A., Nanavati, A. A., and Rajput, N. User-Generated Content Creation and Dissemination in Rural Areas. *ITID 6*, 2 (2010).
8. Chakraborty, D., Medhi, I., Cutrell, E., and Thies, W. Man Versus Machine: Evaluating IVR Versus a Live Operator for Phone Surveys in India. In *DEV* (2013).
9. Desai, S. B., Dubey, A., Joshi, B. L., Sen, M., Shariff, A., and Vanneman, R. Human Development in India: Challenges For a Society in Transition. Tech. rep., Oxford University Press, 2010.
10. D'Silva, K., Marathe, M., Vashistha, A., Borriello, G., and Thies, W. A Mobile Application for Interactive Voice Forums: Design and Pilot Deployment in Rural India. In *DEV* (2014).
11. Eagle, N. txteagle : Mobile Crowdsourcing. *Money* (2009).
12. Grover, A., and Calteaux, K. A voice service for user feedback on school meals. In *DEV* (2012).
13. Gulaid, M., and Vashistha, A. Ila Dhageyso : An Interactive Voice Forum to Foster Transparent Governance in Somaliland. In *ICTD* (2013).
14. Gupta, A., and Thies, W. mClerk: enabling mobile crowdsourcing in developing regions. In *CHI* (2012).
15. Koradia, Z., Aggarwal, P., Seth, A., and Luthra, G. Gurgaon idol: A singing competition over community radio and IVRS. In *DEV* (2013).
16. Lerer, A., Ward, M., and Amarasinghe, S. P. Evaluation of IVR data collection UIs for untrained rural users. In *DEV* (2010).
17. Medhi, I., Patnaik, S., Brunskill, E., Gautama, S. N., Thies, W., and Toyama, K. Designing mobile interfaces for novice and low-literacy users. *TOCHI 18*, 1 (2011).
18. Meeker, M. 2014 Internet Trends. Tech. rep., Kleiner Perkins Caufield Byers, 2014. **http://www.kpcb.com/internet-trends**.
19. Mudliar, P., Donner, J., and Thies, W. Emergent Practices Around CGNet Swara: A Voice Forum for Citizen Journalism in Rural India. *ITID* (2013).
20. Munroe, R. Reddit's new comment sorting system, Oct. 2009. **http://www.redditblog.com/2009/10/reddits-new-comment-sorting-system.html**.
21. Narula, P., Gutheim, P., Rolnitzky, D., Kulkarni, A., and Hartmann, B. MobileWorks : A Mobile Crowdsourcing Platform for Workers at the Bottom of the Pyramid. In *HCOMP* (2011).
22. Ochoa, X., and Duval, E. Quantitative analysis of user-generated content on the web. In *WebEvolve* (2008).
23. Patel, N., Agarwal, S., Rajput, N., Nanavati, A., Dave, P., and Parikh, T. S. A Comparative Study of Speech and Dialed Input Voice Interfaces in Rural India. In *CHI* (2009).
24. Rangaswamy, N., and Cutrell, E. Anthropology, Development and ICTs: Slums, Youth and the Mobile Internet in Urban India. In *ICTD* (2012).
25. Raza, A. A., Rosenfeld, R., Sherwani, J., Milo, C., Alster, G., Saif, U., Pervaiz, M., and Razaq, S. Viral Entertainment as Vehicle for Disseminating Speech Based Services to Low Literate Users. In *ICTD* (2012).
26. Raza, A. A., Ul Haq, F., Tariq, Z., Pervaiz, M., Razaq, S., Saif, U., and Rosenfeld, R. Job opportunities through entertainment: Virally spread speech-based services for low-literate users. In *CHI* (2013).
27. Smyth, T. N., Kumar, S., Medhi, I., and Toyama, K. Where There's a Will There's a Way: Mobile Media Sharing in Urban India. In *CHI* (2010).
28. Su, X., and Khoshgoftaar, T. M. A survey of collaborative filtering techniques. *Adv. in AI* (2009).
29. Vashistha, A., Brady, E., Thies, W., and Cutrell, E. Educational Content Creation and Sharing by Low-Income Visually Impaired People in India. In *DEV* (2014).
30. White, J., Duggirala, M., Srivastava, S., and Kummamuru, K. Designing a Voice-Based Employment Exchange for Rural India. In *ICTD* (2012).