

# Spread and Sustainability:

## The Geography and Economics of Speech-Based Services

Agha Ali Raza, Roni Rosenfeld  
Language Technologies Institute  
School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA  
{araza, roni}@cs.cmu.edu

Farhan ul Haq, Zain Tariq, Umar Saif  
Department of Computer Science  
School of Science and Engineering  
Lahore University of Management Sciences  
Lahore, Pakistan  
{farhan.haq, zain.tariq, umar}@lums.edu.pk

### ABSTRACT

We have been developing techniques for spreading telephone-based services to low-literate people in the developing world, bypassing the need for explicit user training. We achieve this by using entertainment as a viral conduit to spread and popularize development related voice-based services. *Polly*, our telephone-based voice manipulation and forwarding system, has been in continuous operation in Pakistan since May 2012. In this poster, we show the geographical spread of Polly over the initial four months of its deployment. We then describe our attempts at reducing our operating costs by shifting some of them to users, and the impact this had on user behavior, demonstrated via randomized control trials and by the usage of landline vs. mobile phones.

### 1. Introduction

In 2011 Raza et al. [1] launched a telephone-based speech service, called Polly, in Pakistan. Designed for low-literate users, it provided simple entertainment by allowing a caller to record a short message, change it using one of several funny voice modifications, and then forward the manipulated recording to any number of recipients. Seeded with 32 low-skilled office workers in a Pakistani university, Polly spread to more than 2,000 users in just three weeks, generating traffic of more than 10,000 calls. It had to be shut down due to insufficient telephone capacity and funds required to sustain the subsidized service. Later analysis showed that Polly's main use was as a voicemail and group messaging system, and that it was able to spread across gender and age boundaries. However, it did cross class boundaries, remaining mostly among low SES users.

Raza et al. [2] reported the first large-scale sustained deployment of Polly: re-launched in Pakistan in May 2012 with a 30 telephone line capacity, in five months Polly spread to 85,000 users, engaging them in 495,000 interactions, and was continuing its spread to 1,000 new people every day. One major addition to this re-launch was a voice based job search service for low-skilled workers. In its first five months, the job service attracted over 27,000 people, who in turn listened over 279,000 times to job ads and forwarded them over 22,000 times to their friends. [2] also analyzed the traffic, demographics and message contents.

### 2. Geographical Spread

Polly is hosted in Lahore and its seed users were mostly from Okara (a city in Punjab located near Lahore). After initial seeding we made no further effort to publicize Polly. Polly spread quickly via message forwarding and word-of-mouth. Within the first four months of its operation it has received calls

from, and delivered messages to, all four provinces of Pakistan, as well as to Belgium (19), UAE (11), Saudi Arabia (4), Spain (4), Oman (3), India (1) and Luxembourg (1). Figure 1 shows some snapshots of the domestic geographical spread of Polly.

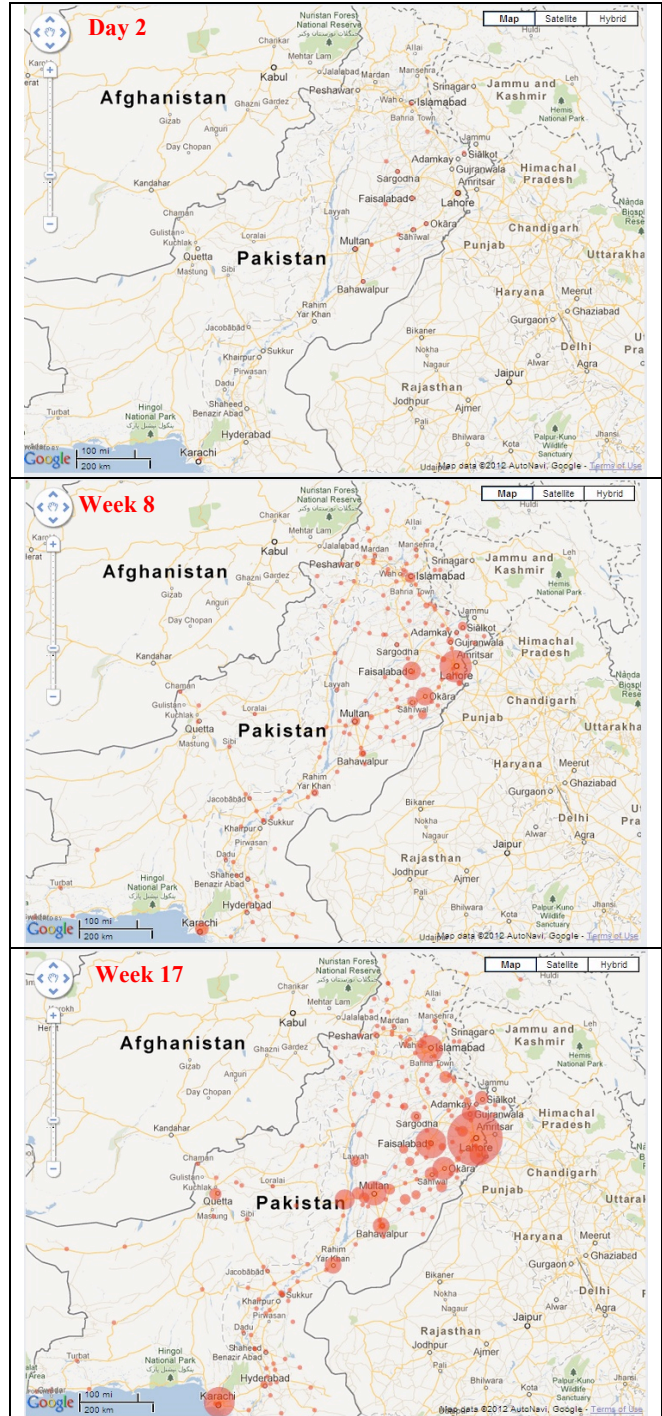


Figure 1: Geographical spread of Polly (only Pakistan shown)

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

### 3. Cost and Sustainability

Currently Polly completely resides on a single server that is hosted by a local telecom provider in Pakistan. This local hosting has reduced our outgoing call airtime cost to \$0.023/minute. However, we had to look for additional ways of cutting our costs due to the high call volume, peaking at 7,000 calls per day, and averaging around 3 minutes.

[2] describes how the number of daily toll-free (“missed call” based) calls to Polly were gradually restricted for each user. Once users exceeded their daily quota, they were invited to a caller-paid phone number, which provided a similar (in fact, slightly enhanced) Polly experience. Although more than 4,500 users used the caller-paid line to make more than 12,000 calls, the majority of users chose to avoid the call charges by waiting until the following day. For more details, see [2].

After limiting toll-free calls to one per user per day, we were still paying for 4,000 calls/day, costing nearly \$300/day in airtime alone. We did not want to eliminate toll-free calls altogether, because we did not want to lose the lower-SES users, who are our primary target. Upon further analysis, we found that of the 4,000 Polly-paid daily calls, some 60% were interactions initiated by users, and 40% were *message deliveries* – calls initiated by Polly to deliver previously scheduled messages to their recipients. To further reduce costs, we focused on message delivery calls. We wanted to test how willing people will be to pay call charges in order to retrieve messages sent to them by friends. We therefore replaced the original model with one where the recipient is just notified that a message has been recorded for him and that he may call our caller-paid line to listen to it. The caller-paid line was then modified to add the following functionality:

- a. When an experienced user calls, if they have any messages waiting Polly informs them of their number and prompts them to choose between listening to them and going directly to interacting with Polly.
- b. When an inexperienced user calls, if they have any messages waiting Polly informs them of their number and starts playing them one-by-one. After each message it prompts the user to choose between continuing to listen to messages and switching to normal interaction with Polly.

A user is considered experienced if he has interacted with Polly at least three times before. Messages that have been skipped are retained for future retrieval. Job-related calls were not affected and were always delivered at our expense.

We send message notifications using SMS and optionally short voice follow-up calls. We conducted this experiment in three phases:

#### 1. SMS-only notification

Under this experiment, an SMS was sent to the intended recipient of a message immediately upon its scheduling, briefly informing him that a message has been recorded for him, listing the sender’s phone number and Polly’s caller-paid number for retrieval. A total of 7,893 such SMSs were sent during this stage of the experiment, which resulted in 514 (merely 6.5%) caller-paid retrieval calls. Of note, 57% of these calls were made within one hour of the SMS delivery, and 85% were made within 24 hours.

#### 2. SMS notification with a 24-hour voice follow-up

We conducted this as a two arm randomized controlled trial. Out of all the SMS notifications that did not result in retrievals within 24 hours, half were randomly chosen for sending voice-

call reminders. Polly called each of these intended recipients and briefly announced that a message is waiting for them, playing the sender’s name in their own recorded voice, and explaining that, to listen to the messages, the receiver needs to call back the number from which they were just called (which was the caller-paid line). The main motivation for voice call reminders was to help recipients who may not have understood the SMS, perhaps due to lack of literacy. Also, with voice alerts the user was not required to type any numbers to access their messages, but could simply call back using their mobile phone’s history. During the period of this trial, 4.7% of the delivered SMS alerts triggered retrievals within 24 hours, and hence were not assigned to any arm. Of the remaining messages, 2,693 were randomly assigned to SMS only (the control arm), out of which 26 (0.96%) were eventually retrieved. We attempted voice reminders for the other 2,681 messages, resulting in 113 retrievals. So, the message pickup rate of the treatment group was 4.2% vs. 0.96% for the control group, with an estimated effect size of  $4.2\% - 0.96\% = 3.24\%$ , a relative increase of 57% in the retrieval rate.

Of note, among those who responded to the voice calls by retrieving their message, new users were somewhat over-represented (40% vs. their 32% share of all recipients), suggesting that unfamiliarity and lack of reading/comprehension skills are a factor in failing to retrieve messages.

#### 3. SMS notification with a 1 hour voice follow-up

This was also conducted as a two arm RCT, very similar to the previous one, except that voice call reminders were sent 1 hour after SMS delivery. During the period of this trial, 3.5% of the delivered SMS alerts triggered retrievals within one hour, and hence were not assigned to any arm. Of the remaining messages, 868 were randomly assigned to the SMS-only (control) arm, out of which 21 (2.4%) were subsequently retrieved. We attempted voice reminders for the other 880 messages, resulting in 61 retrievals. The message pickup rate of the treatment group was therefore  $6.9\% - 2.4\% = 4.5\%$ , a relative increase of 63% in the retrieval rate.

#### 4. Landline vs. Mobile phone usage

In Pakistan, telephone numbers readily reveal whether they belong to a landline or to a mobile phone. Polly’s service is provided via a landline, making it cheaper to call it from a landline than from a mobile phone.

Only 0.32% of the more than 400,000 toll-free calls were made from landlines. In contrast, landlines accounted for almost 4% of the caller-paid calls. This demonstrates some cost sensitivity on the part of users when they need to pay for the call. But this is only a lower bound because, unfortunately, we do not know how many of the users who made called-paid calls had access to a landline.

### 5. REFERENCES

[1] A. Raza, C. Milo, G. Alster, J. Sherwani, M. Pervaiz, S. Razaq, U. Saif, and R. Rosenfeld, “Viral entertainment as a vehicle for disseminating speech-based services to low-literate users,” in *International Conference on Information and Communication Technologies and Development (ICTD)*, vol. 2, 2012.

[2] A. Raza, F. Ul Haq, Z. Tariq, M. Pervaiz, S. Razaq, U. Saif, and R. Rosenfeld, “Come for the fun, stay for the jobs: Virally spreading speech-based services for low-literate users,” *Submitted for Publication*.