Project Polly

Develop *Viral Entertainment* as a vehicle for disseminating *Development* related telephone based services

1. **Introduce** and **popularize** speech interfaces
2. Use Entertainment as a **Viral Conduit** for delivering **core development services**
3. Setup an **Experimental Testbed** for testing speech interface choices
4. Provide **Entertainment**

Polly

Polly is a **telephone-based, voice-based** application which allows users to make a short recording of their voice, **modify it** and **send the modified version to friends**.

Carnegie Mellon University

2012-13 Large Scale Deployment

- Launched on May 09, 2012 in Lahore, Pakistan:
  - Seeded with **5 low-literate people**
- After **1** year:
  - **636,000** calls
  - **165,000** users
  - At its peak it was spreading to **1,000 new people daily**

  - **34,000** people used the job search service
  - listened **386,199** times to **728** job ads
  - and **19,000** users forwarded them **34,000** times to their friends.

This Talk: Analysis of User Behavior

1. Do users fall into naturally distinct groups if they are clustered based on their activity profile?
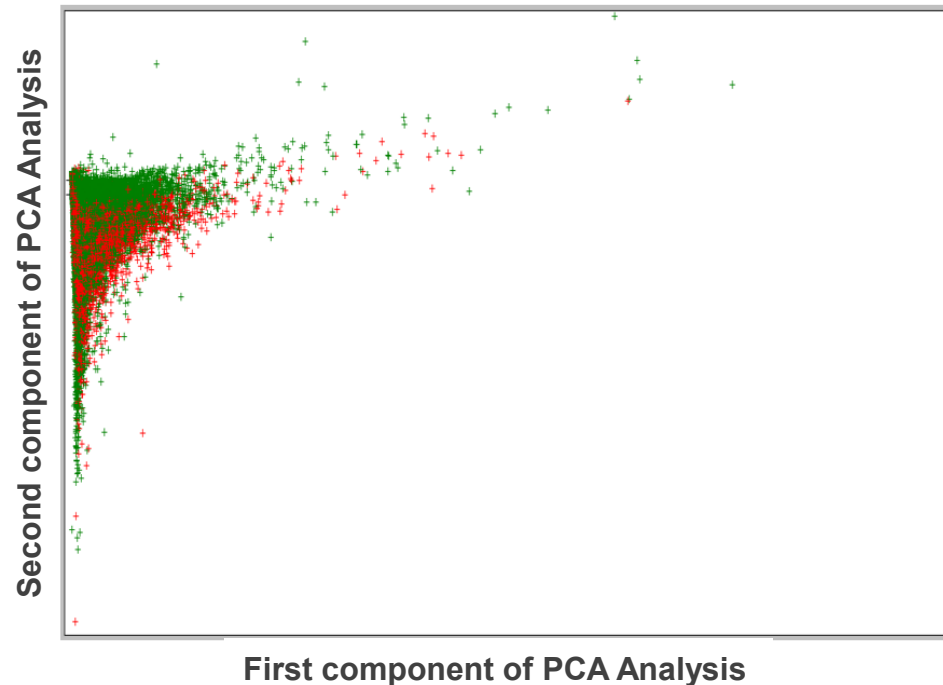
Does more experience using Polly lead to:

2. an improvement in users' interaction skills?
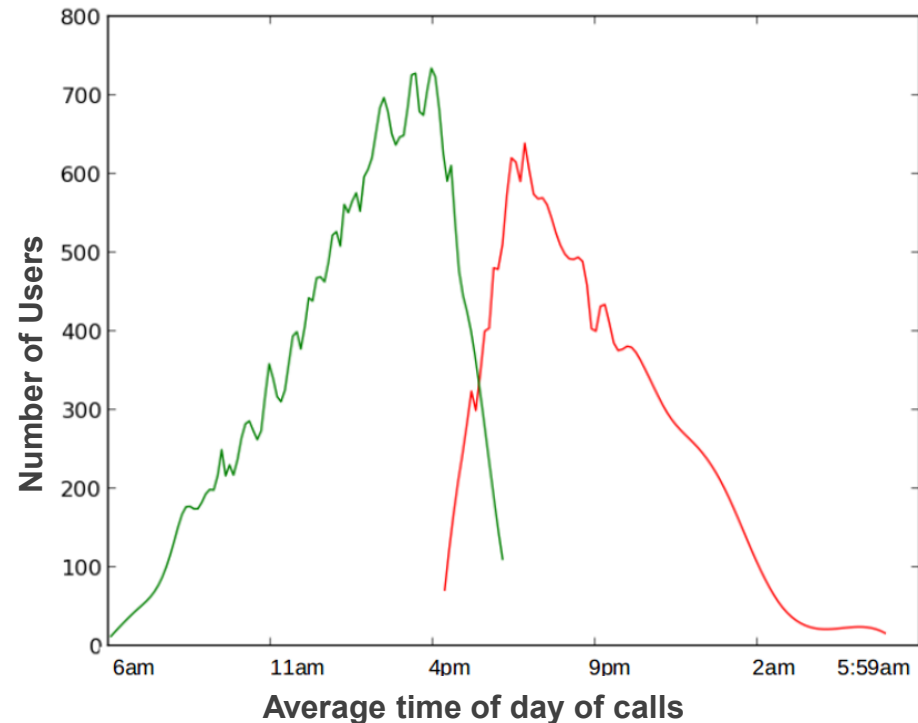
3. any change in usage preferences?

User Clustering

- **63,023** users who had at least one active interaction with the system.

- Each user is represented by features based on:
  - types of calls and initial interaction
  - perseverance
  - preferred time-of-day
  - choices made during the calls
  - pattern of social connectivity

User Clustering

Second component of PCA Analysis (y-axis)

First component of PCA Analysis (x-axis)

Number of Users (y-axis)

Average time of day of calls (x-axis)

- 63,023 active users
- 2 clusters

- Result: No clear separation

Results:
Most informative feature: avg. time of day of calls
- Mid-day vs. evening users
- Midday users are more active, more connected.

Carnegie Mellon University

Improvement in users' interaction skill

We focus on users' interactions with the first (main) menu of Polly.

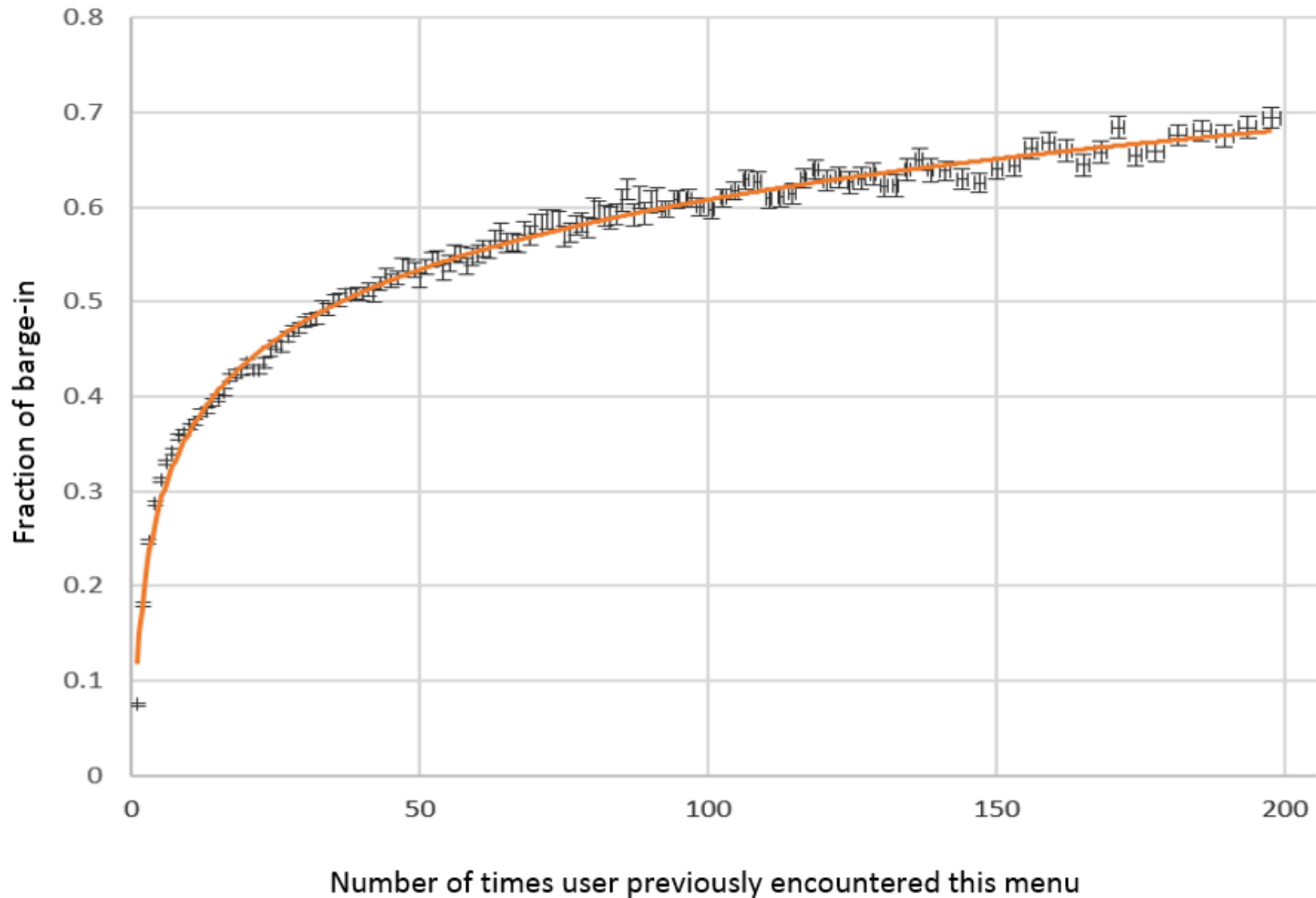- There may be none to several menu-interactions within one call.

We track the prevalence of:

1. Barge-in (pressing button before end of instructions)

2. Invalid button presses
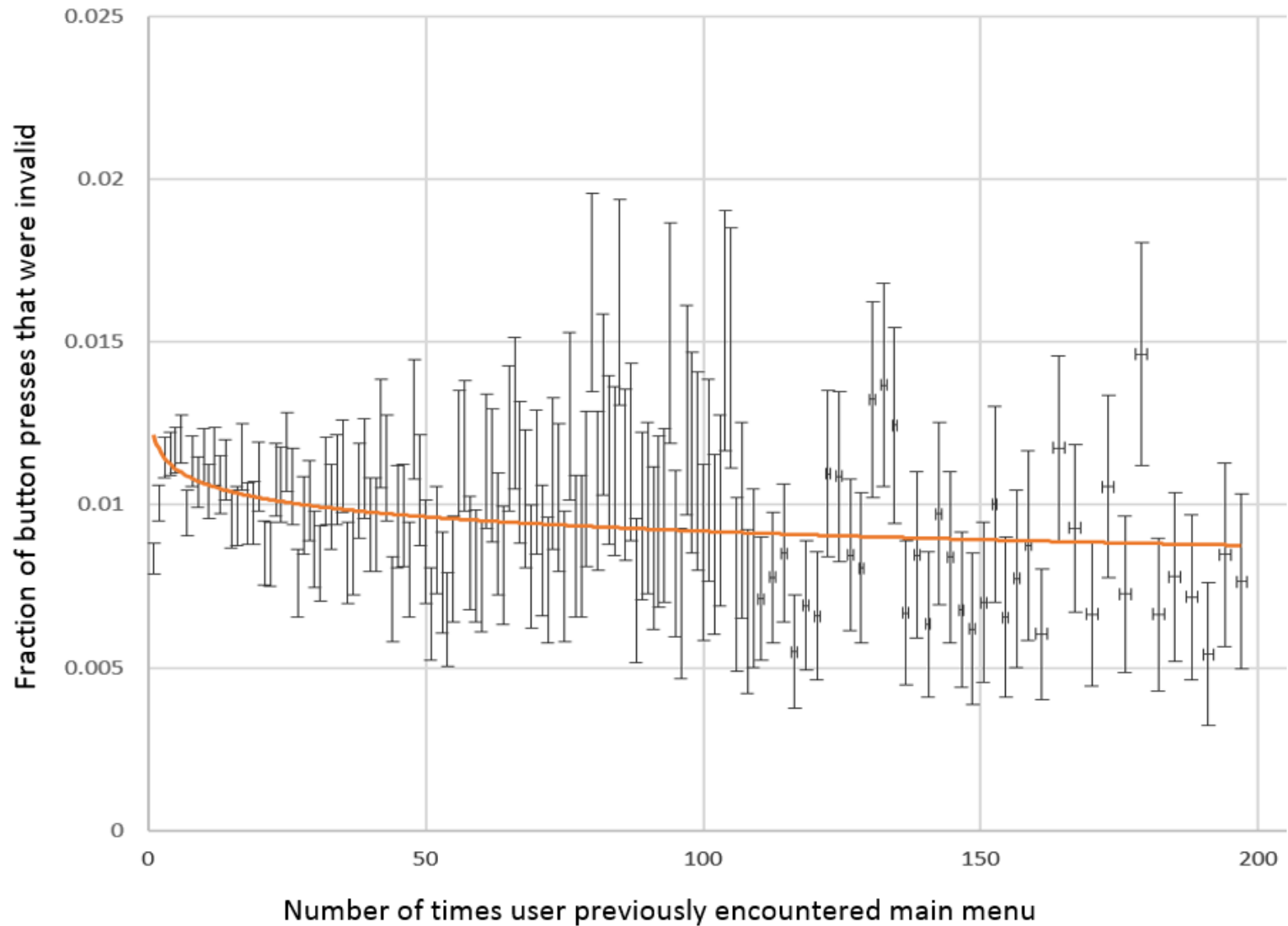
3. Unsuccessful forwarding attempts.

50,414 users, **292,951** calls, **934,742** menu interactions.
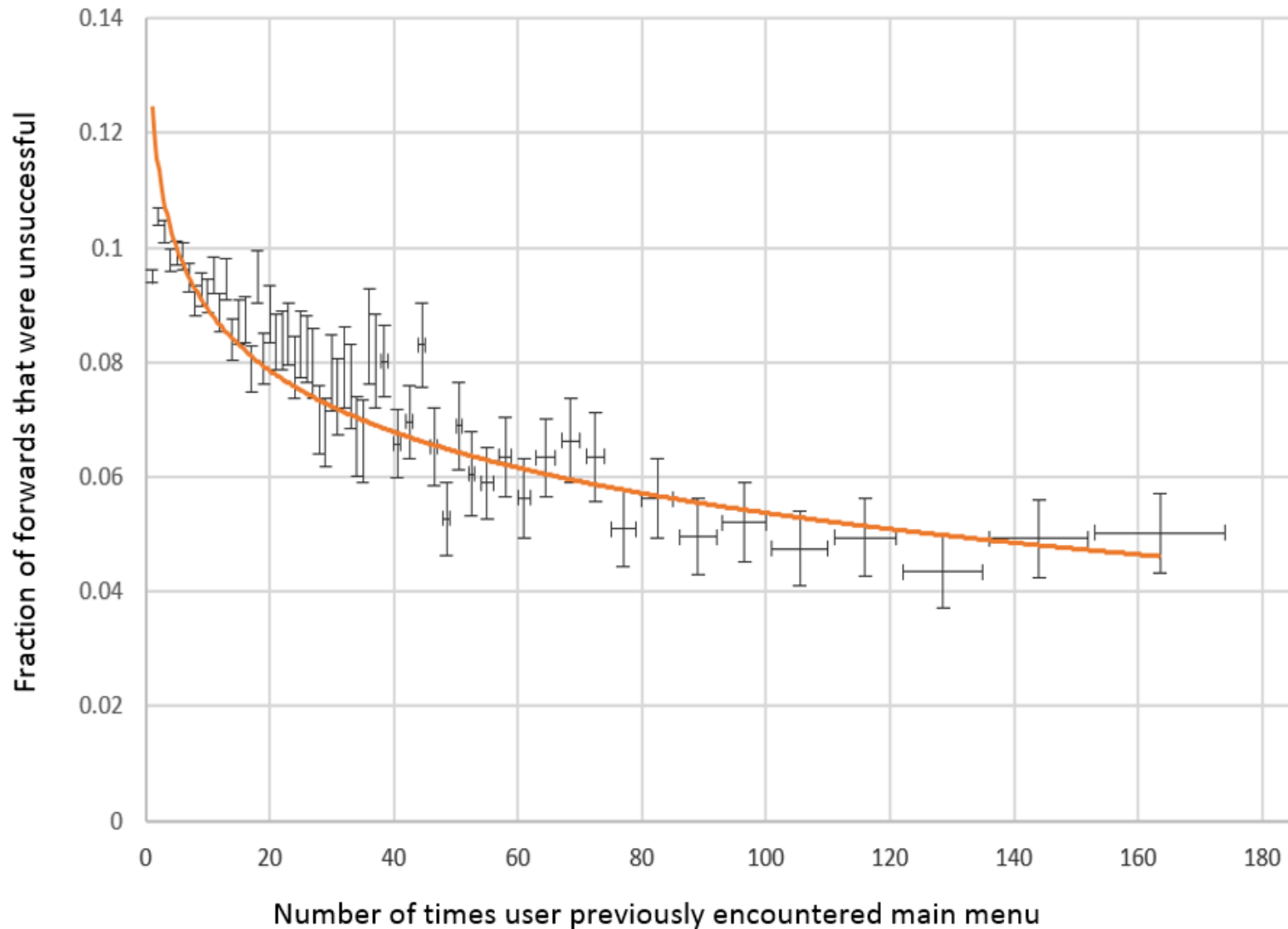
Carnegie Mellon University

Barge-in Behavior

Invalid Button Pressed

Unsuccessful Forwarding attempts

Caveat: Confounding User Types

- This analysis lumps together long-term and short-term users.

- Hence, it confounds true learning by any one user with differences between the different user types.

  - Long-term users may be more adept at using IVR systems to start with.

Carnegie Mellon University

Changes in User Preferences:
(now controlling for user type)

We define 3 user sets and explore changes in their usage patterns as a function of their *experience* (number of prior Polly calls):

- ***Short-term users:*** 2,701 users who interacted with Polly <u>exactly</u> 5 times.

- ***Intermediate-term users:*** 1,862 users who interacted with Polly <u>exactly</u> 10 or 11 times.

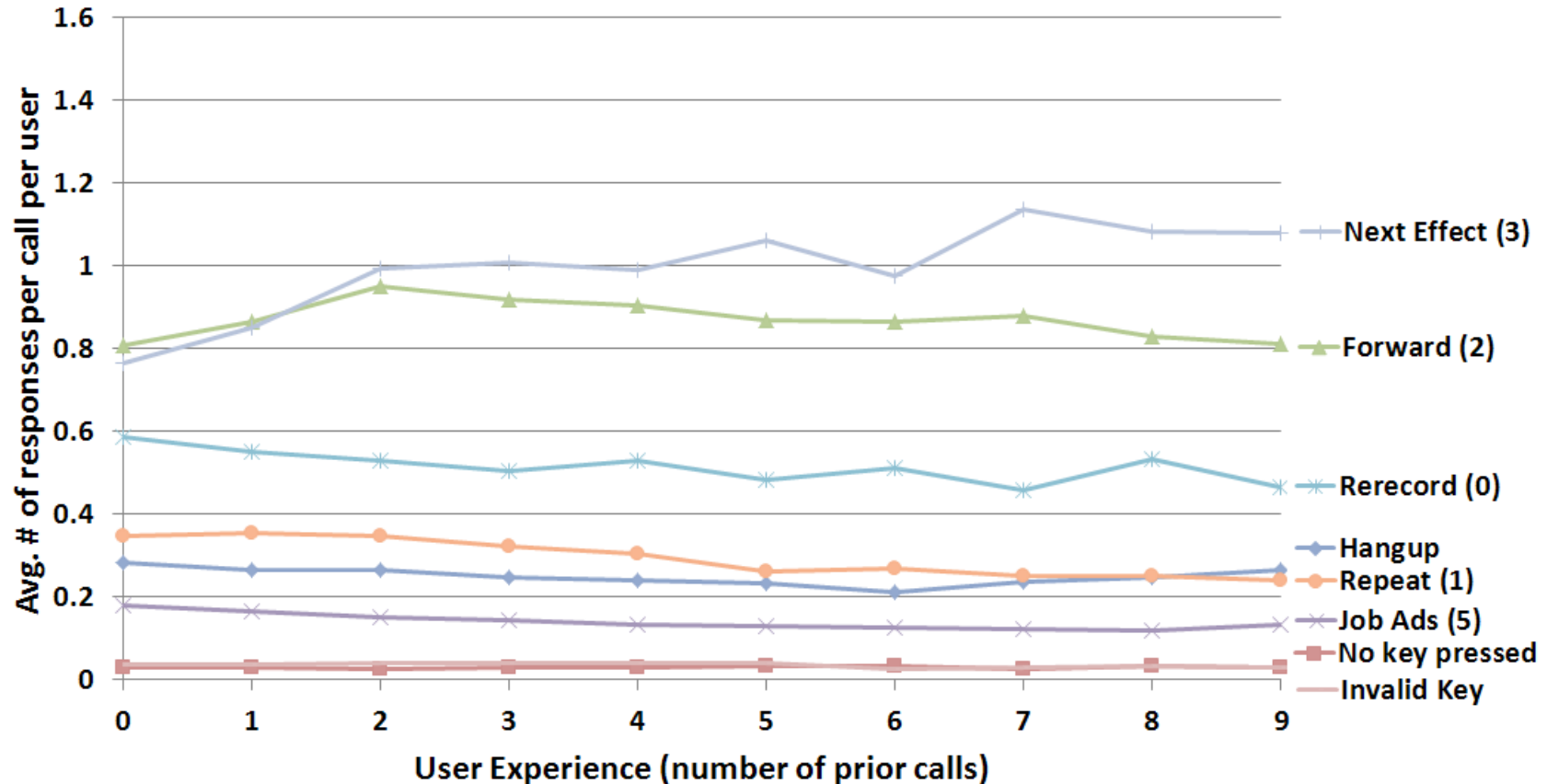- ***Long-term users:*** 1,523 users who interacted with Polly 30+ times.

Menu Choices vs. Experience

Short-term Users (made exactly 5 calls each)
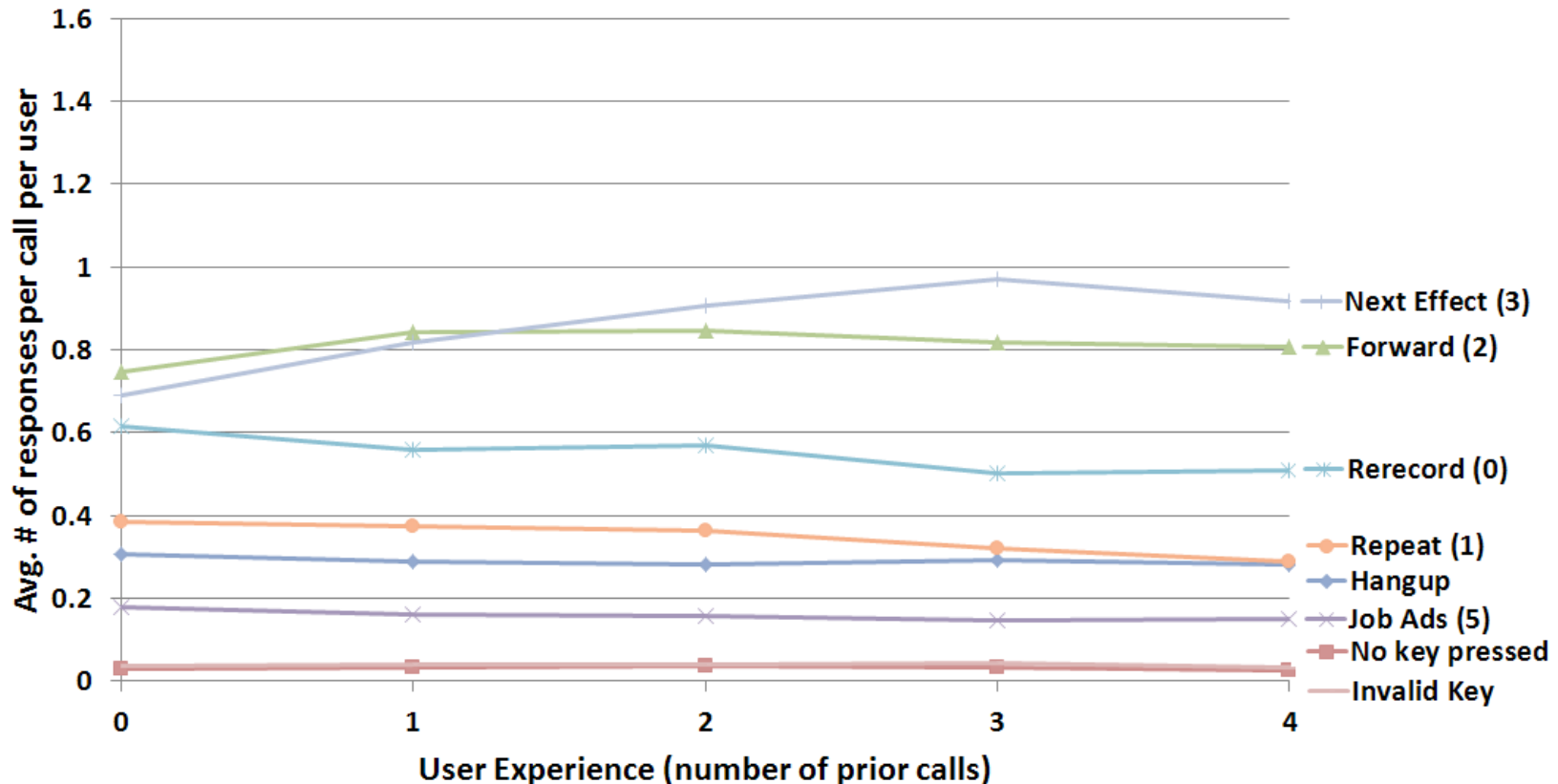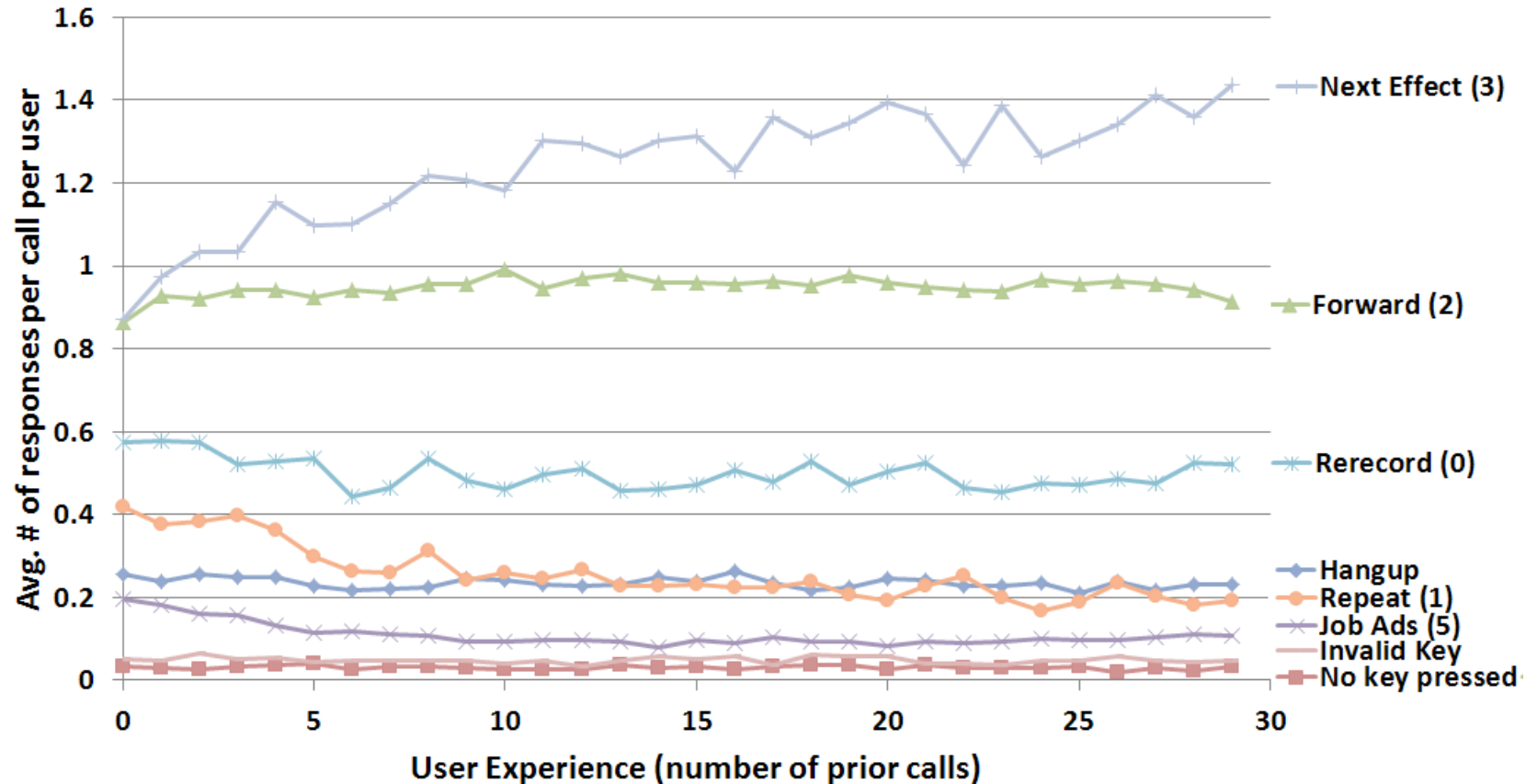
Menu Choices vs. Experience

Intermediate-term Users (made exactly 10 or 11 calls)

Compared to short-term users; activity starts at a higher level and climbs higher!

Menu Choices vs. Experience

Long-term Users (made 30+ calls)

Compared to intermediate-term users; activity starts at a higher level and climbs higher!

Common Trends among all user sets

- The Tendency to:

 1. Press 2 (forward) starts off at a high value and stabilizes

 - "Back Channel"
 - System's popular image as a messaging system

 2. Press 3 (next effect) increases with experience

 - System Exploration
 - Hunting for an effect of choice

 3. Press 0 (re-record) and 1 (repeat) decreases with experience

 - Adapting to 15 seconds recording limit
 - Playing with the system, alone or with friends

- Invalid button presses are rare

Early Differences in Call Complexity
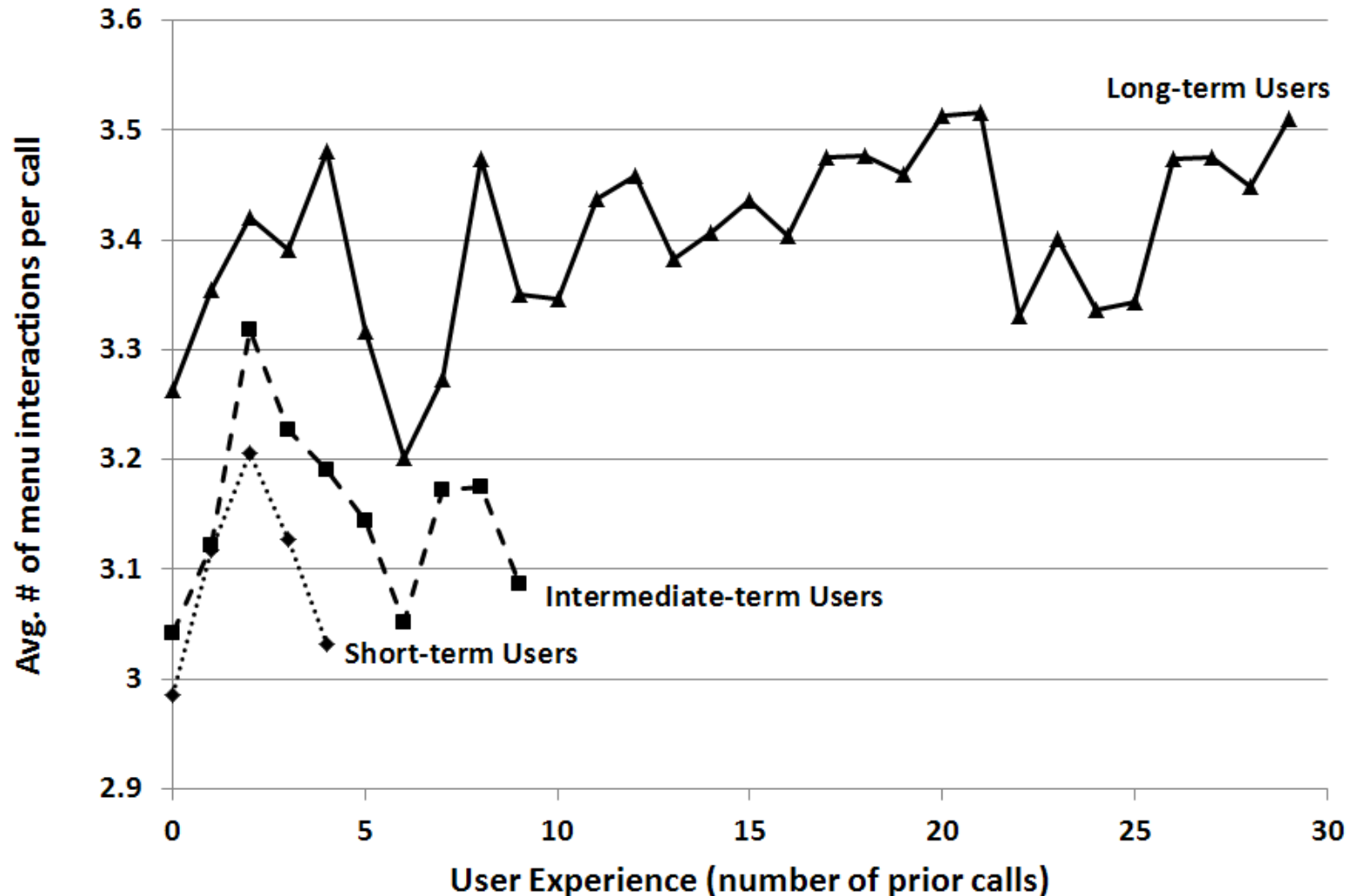
Can be used to predict user retention!

Main Findings

1. Users don't fall into neat groups. Rather, form a continuum.
Significant evidence for learning:

- Increased use of barge-in
- Decreased frequency of failed forwards

2. Significant difference in usage choices between long-term users and (short, intermediate)-term users.

3. Early differences in call complexity among the different user types

Carnegie Mellon University

Current Work

- Re-launch Polly in Pakistan to serve as a clearinghouse for jobs and skill-training opportunities

- Make Polly viral in India with better mechanisms of populating Job Ads and tracking employment outcomes

- Much more analysis to be done
  - Detailed logs, audio recordings,...
  - Collaborators welcome!

Thank you!

Behavior Analysis of Low-literate Users of a Viral Speech-based Telephone Service

Haohan Wang, Agha Ali Raza, Yibin Lin, Roni Rosenfeld
Language Technologies Institute, Carnegie Mellon University

Goal: Using *Viral Entertainment* as a vehicle for disseminating *Development* related telephone based services

Polly: A **telephone-based, voice-based** entertainment service that became viral among low-literate users in Pakistan

- Seeded with **5**
- It reached **165,000 users**, who took part in **636,000 calls**

This talk

1. Do users fall into naturally distinct groups if they are clustered based on their activity profile?

Does more experience using Polly lead to:

2. an improvement in users' interaction skills?

3. any change in usage preferences?