Skill Squatting Attacks on Amazon Alexa

Deepak Kumar, University of Illinois

Riccardo Paccagnella, University of Illinois

Paul Murley, University of Illinois

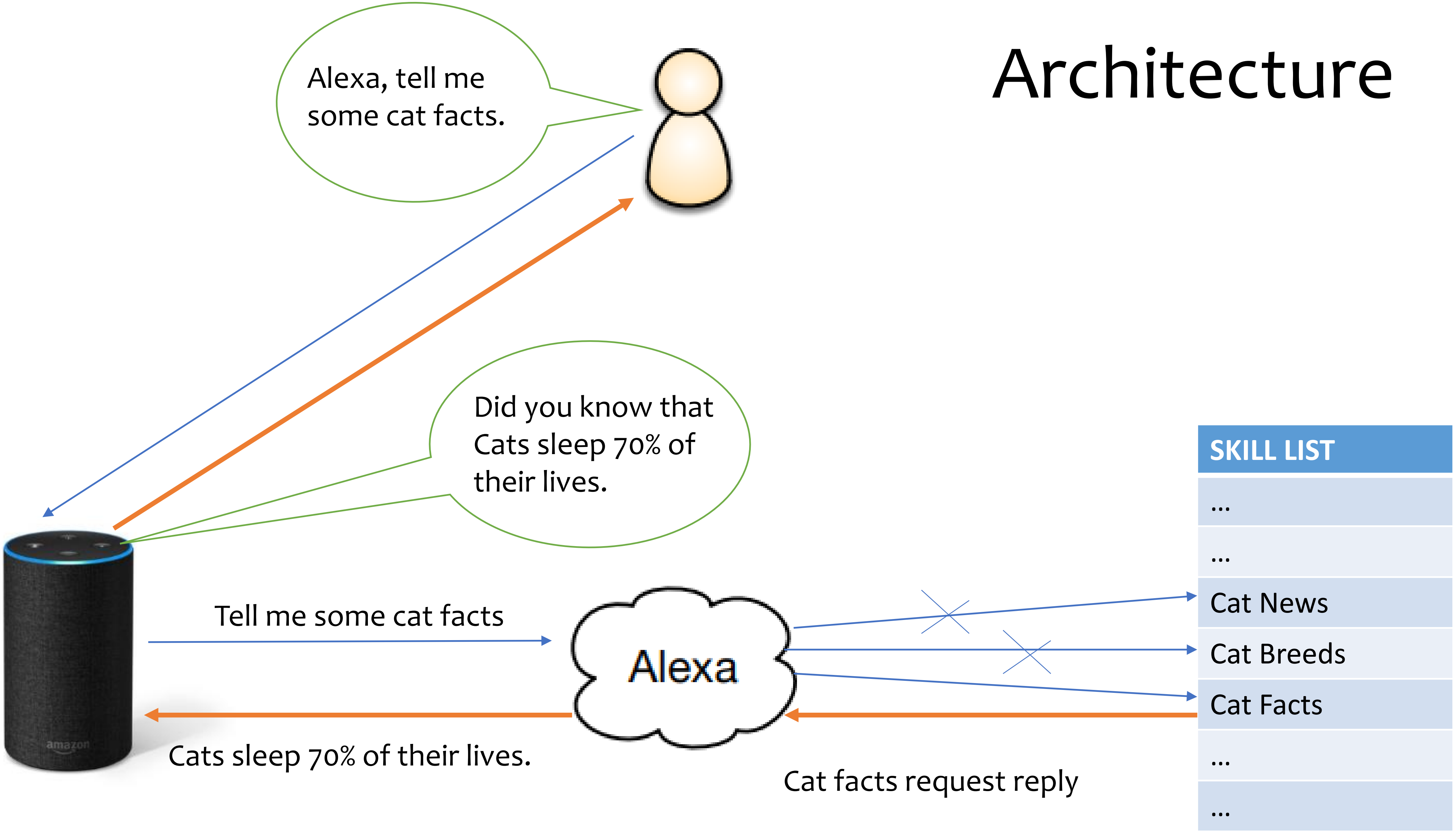
Eric Hennenfent, University of Illinois

Joshua Mason, University of Illinois

Adam Bates, University of Illinois

Michael Bailey, University of Illinois

Architecture



Alexa, tell me some cat facts.

Did you know that Cats sleep 70% of their lives.

Tell me some cat facts

Alexa

Cats sleep 70% of their lives.

Cat facts request reply

SKILL LIST
…
…
Cat News
Cat Breeds
Cat Facts
…
…

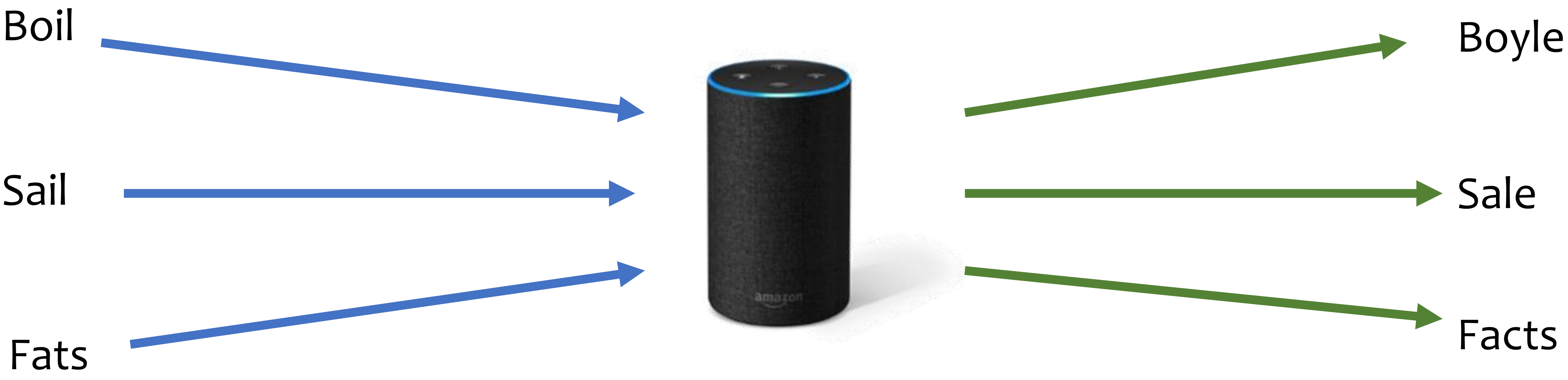
Skill Squatting Attacks

Can Alexa errors be leveraged to cause harm to end users?

An attacker can leverage predictable errors in Alexa to route users to skills that they didn't intend to go to

How can you tell which errors are predictable?

Send speech samples to Alexa, figure out where it goes wrong.



Boil

Sail

Fats

Boyle

Sale

Facts

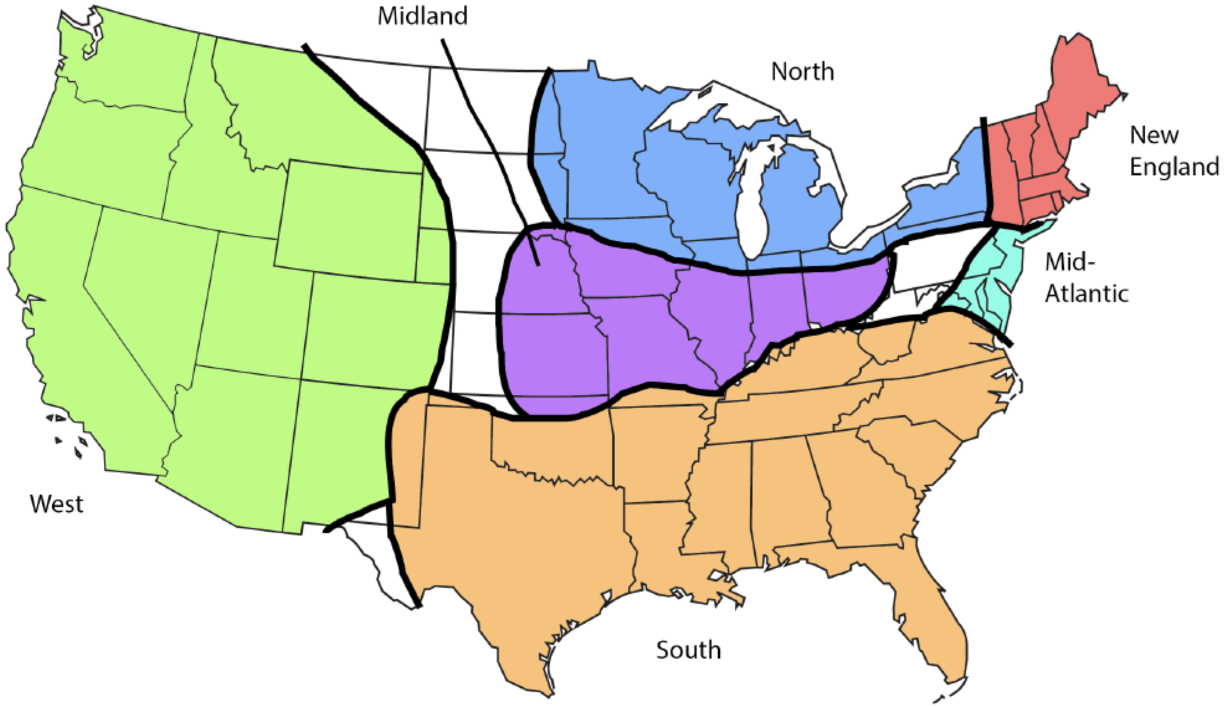
Speech Corpus

NSP The Nationwide Speech Project (NSP)

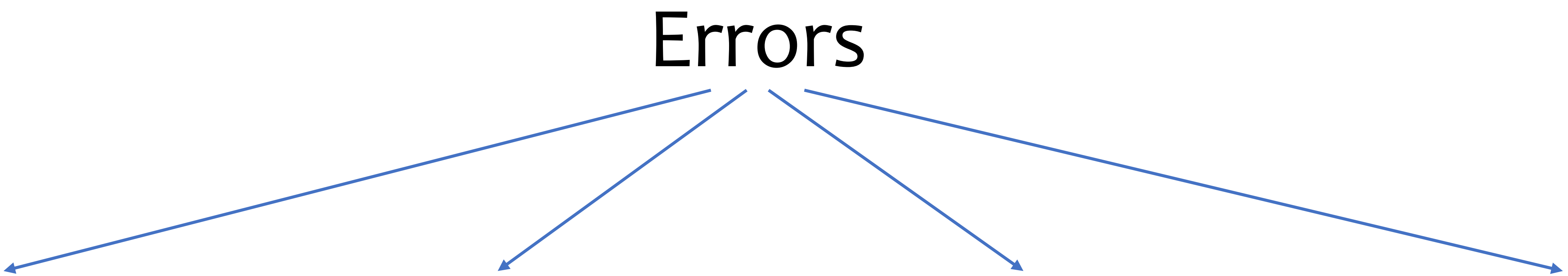
60 speakers from six geographical “dialect-regions” Mid-Atlantic, Midland, New England,
North, South, and West—over a map of the United States.

Five male and five female speakers
from each region provide a set of 188 single-word recordings = (11,460 audio samples)

Send each speech sample to Alexa 50 times, providing us 573,000 transcriptions across the 60 speakers



24 words for which Alexa consistently outputs one wrong interpretation.



Errors

Homophones

Have the same pronunciation

“Sail” – “Sale”

“Sell” – “Cell”

Compound Words

Alexa appears to break these into their constituent words.

“superhighway” - “super highway”

“outdoors” - “out doors”

Phonetic Confusion

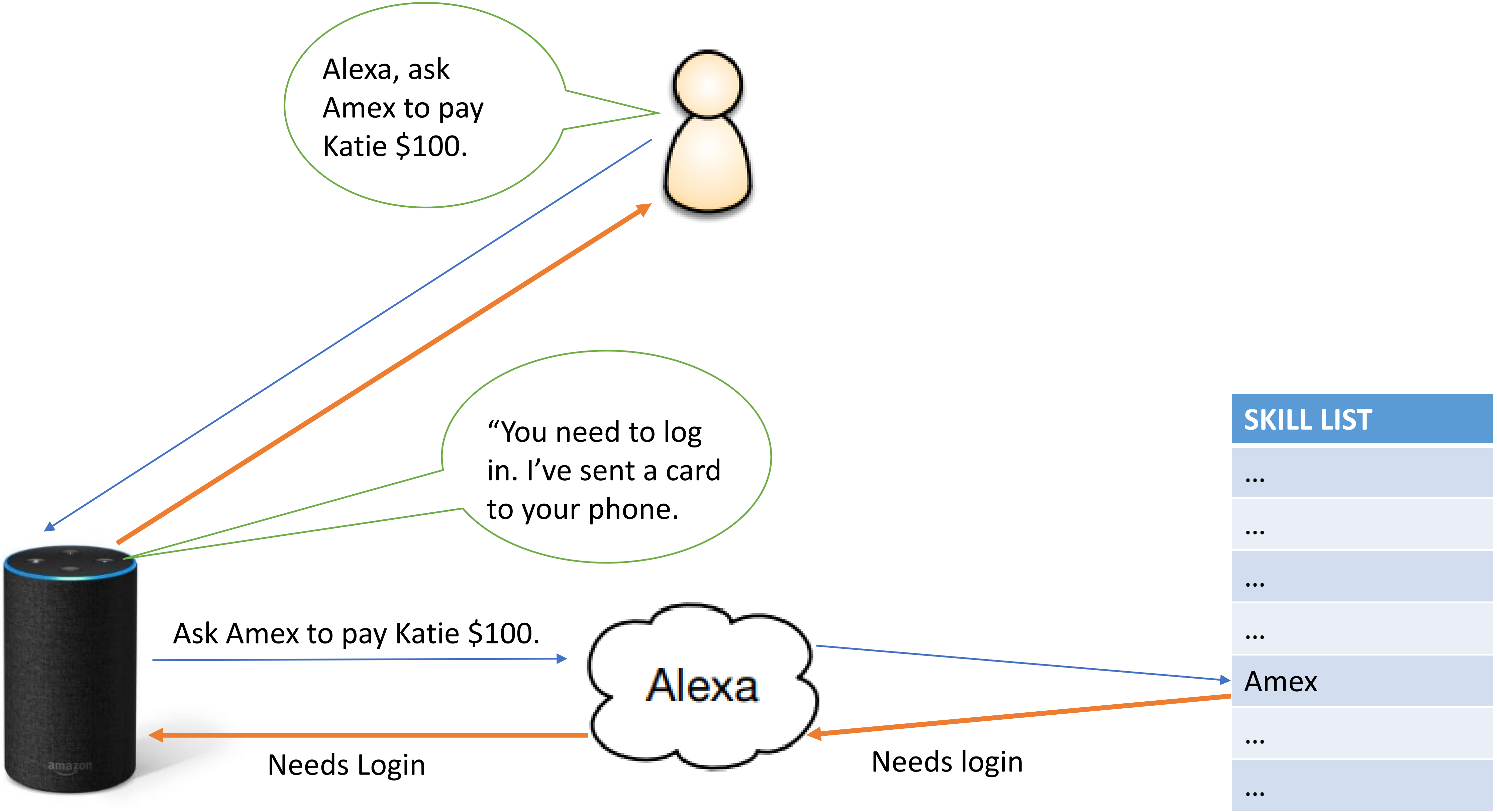
Alexa often misunderstands certain specific phonemes within words while correctly interpreting the rest of them.

“Dull” (D AH L) –
“Doll” (D AA L)

Other Errors

Not easily explainable

“mill” to “no”,
“full” to “four”
“earthy” to “Fi”.



Alexa, ask Amex to pay Katie $100.

“You need to log in. I’ve sent a card to your phone.

Ask Amex to pay Katie $100.

Alexa

Needs Login

Needs login

SKILL LIST

…

…

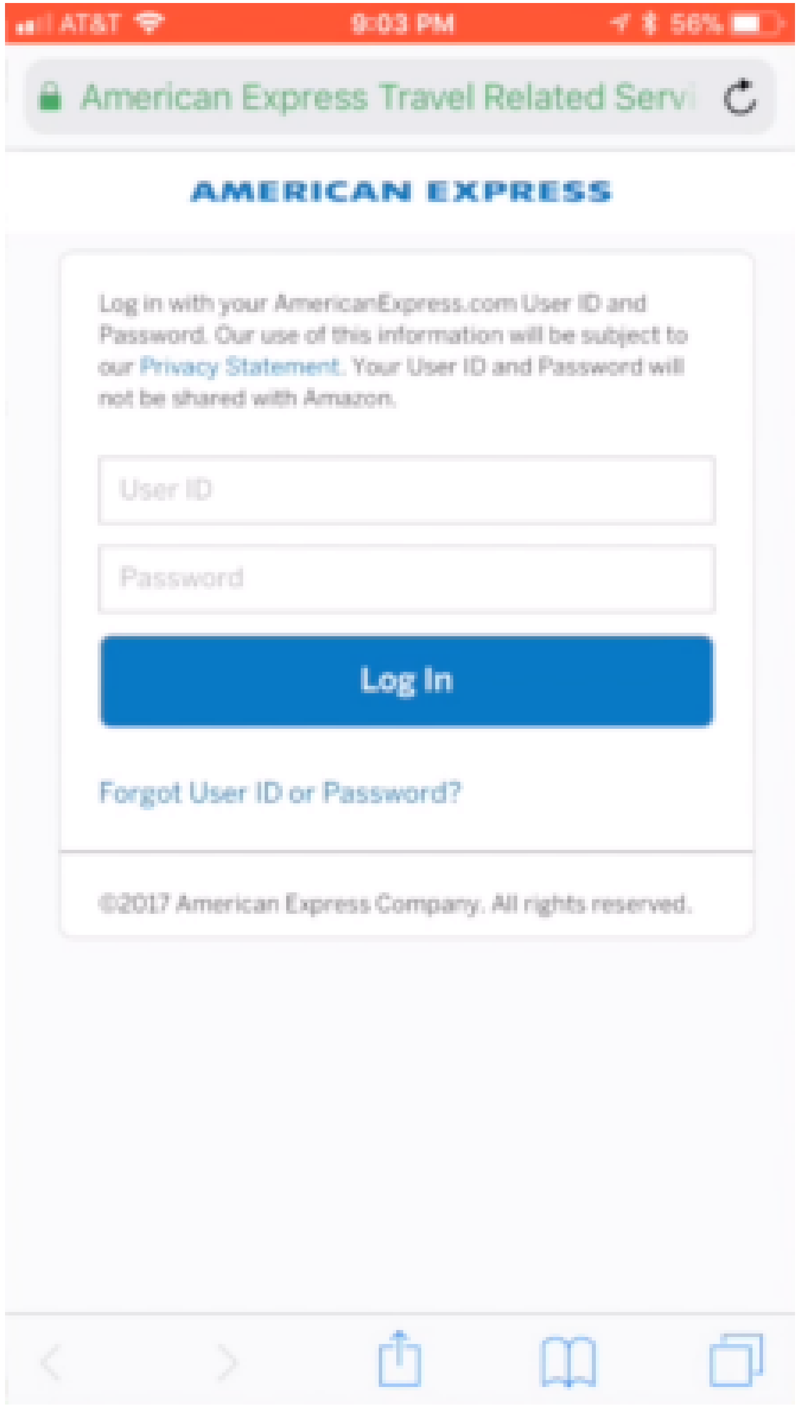
…

…

Amex

…

…

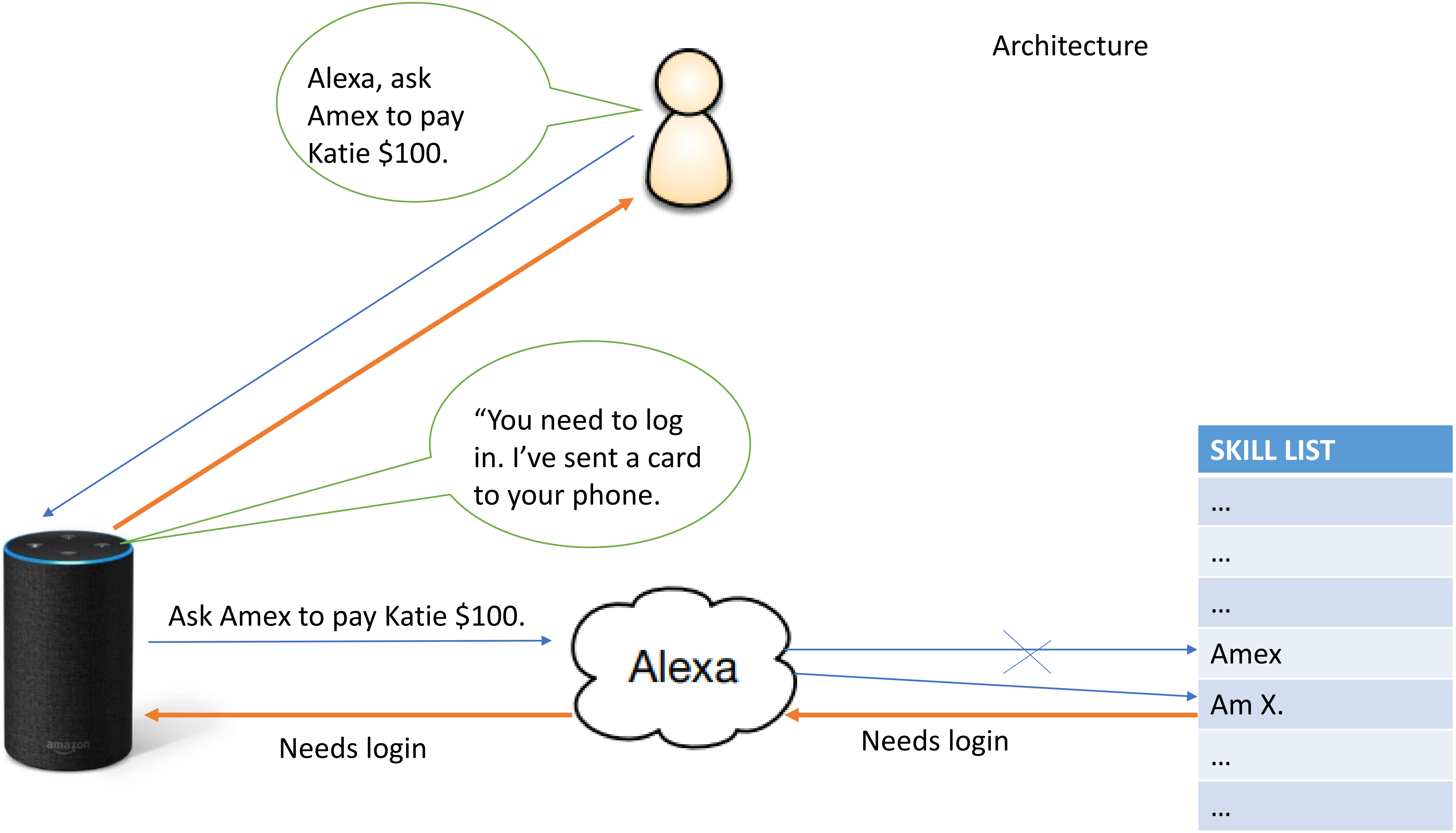


Log in with your AmericanExpress.com User ID and Password. Our use of this information will be subject to our Privacy Statement. Your User ID and Password will not be shared with Amazon.

Forgot User ID or Password?

©2017 American Express Company. All rights reserved.

Architecture



Alexa, ask Amex to pay Katie $100.

“You need to log in. I’ve sent a card to your phone.

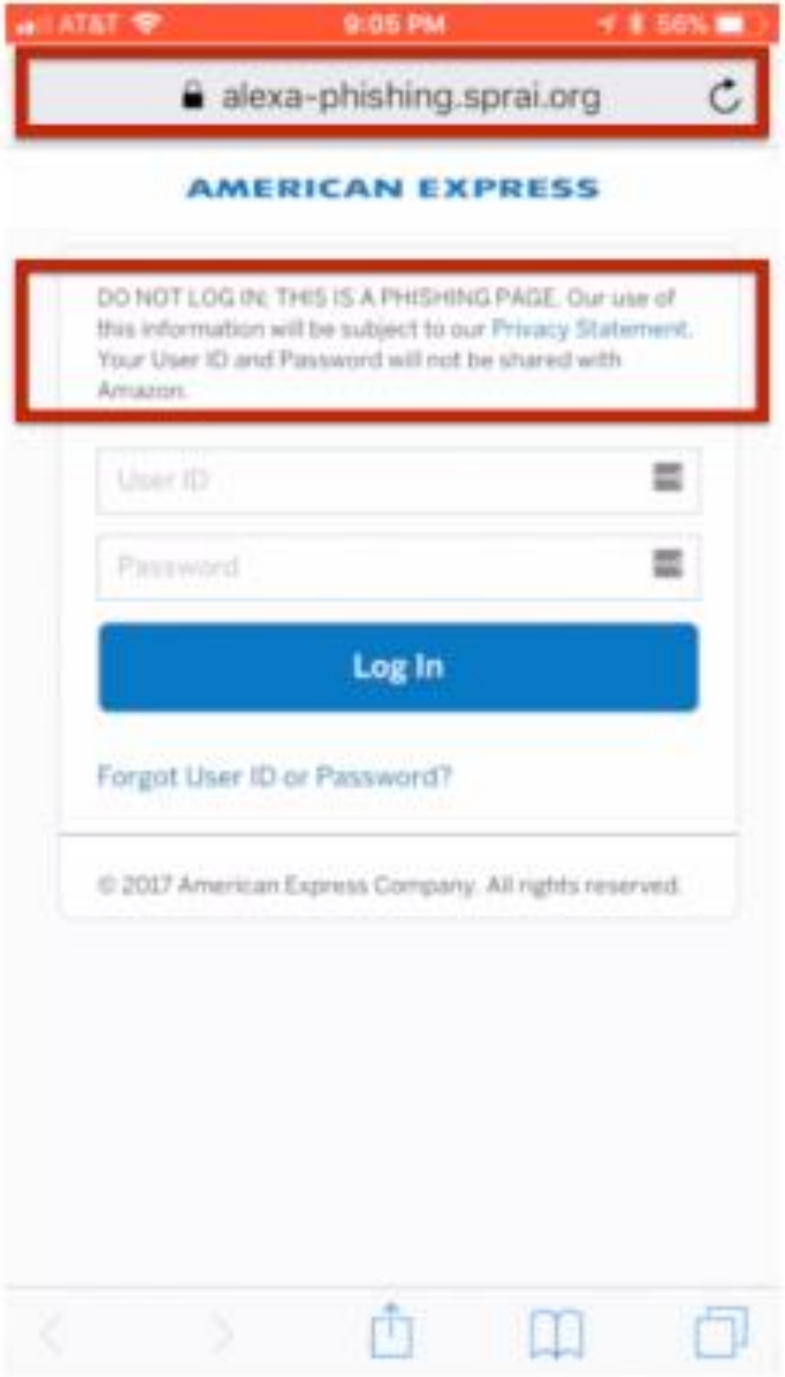
Ask Amex to pay Katie $100.

Alexa

Needs login

Needs login

SKILL LIST
…
…
…
Amex
Am X.
…
…



alexa-phishing.sprai.org

AMERICAN EXPRESS

DO NOT LOG IN; THIS IS A PHISHING PAGE. Our use of this information will be subject to our Privacy Statement. Your User ID and Password will not be shared with Amazon.

© 2017 American Express Company. All rights reserved.

Test Process

Skills Created – 135…….90

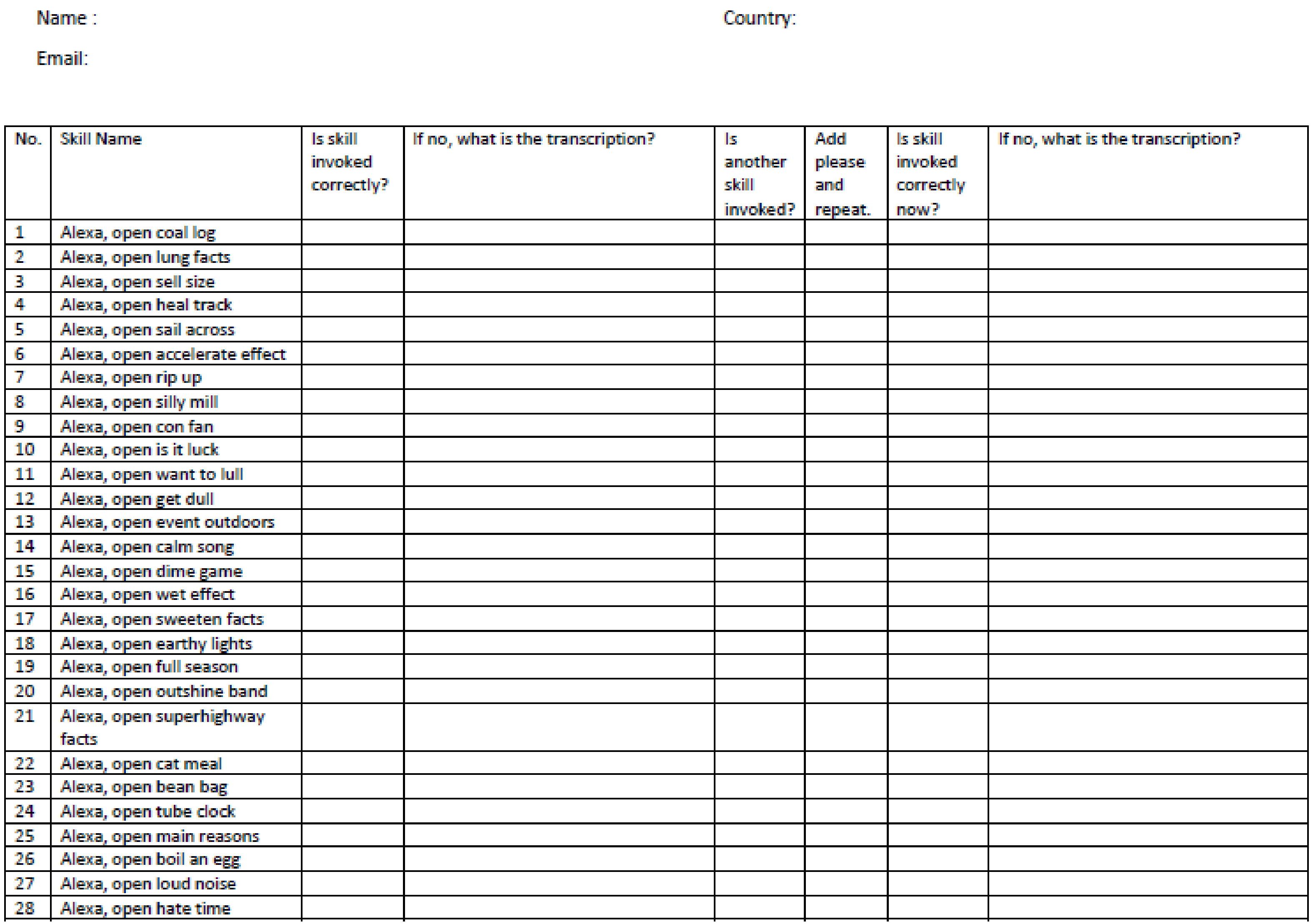
Countries – China, India, Bangladesh, Iran, USA, Africa

Male – 1, Female-1

Assumptions:

1. Alexa is a black box.
2. Triggering skills in a development environment is same as that to triggering publicly available skills.

User Survey Form



Name :

Country:

Email:

No.	Skill Name	Is skill invoked correctly?	If no, what is the transcription?	Is another skill invoked?	Add please and repeat.	Is skill invoked correctly now?	If no, what is the transcription?
1	Alexa, open coal log						
2	Alexa, open lung facts						
3	Alexa, open sell size						
4	Alexa, open heal track						
5	Alexa, open sail across						
6	Alexa, open accelerate effect						
7	Alexa, open rip up						
8	Alexa, open silly mill						
9	Alexa, open con fan						
10	Alexa, open is it luck						
11	Alexa, open want to lull						
12	Alexa, open get dull						
13	Alexa, open event outdoors						
14	Alexa, open calm song						
15	Alexa, open dime game						
16	Alexa, open wet effect						
17	Alexa, open sweeten facts						
18	Alexa, open earthy lights						
19	Alexa, open full season						
20	Alexa, open outshine band						
21	Alexa, open superhighway facts						
22	Alexa, open cat meal						
23	Alexa, open bean bag						
24	Alexa, open tube clock						
25	Alexa, open main reasons						
26	Alexa, open boil an egg						
27	Alexa, open loud noise						
28	Alexa, open hate time						

Predictable Errors

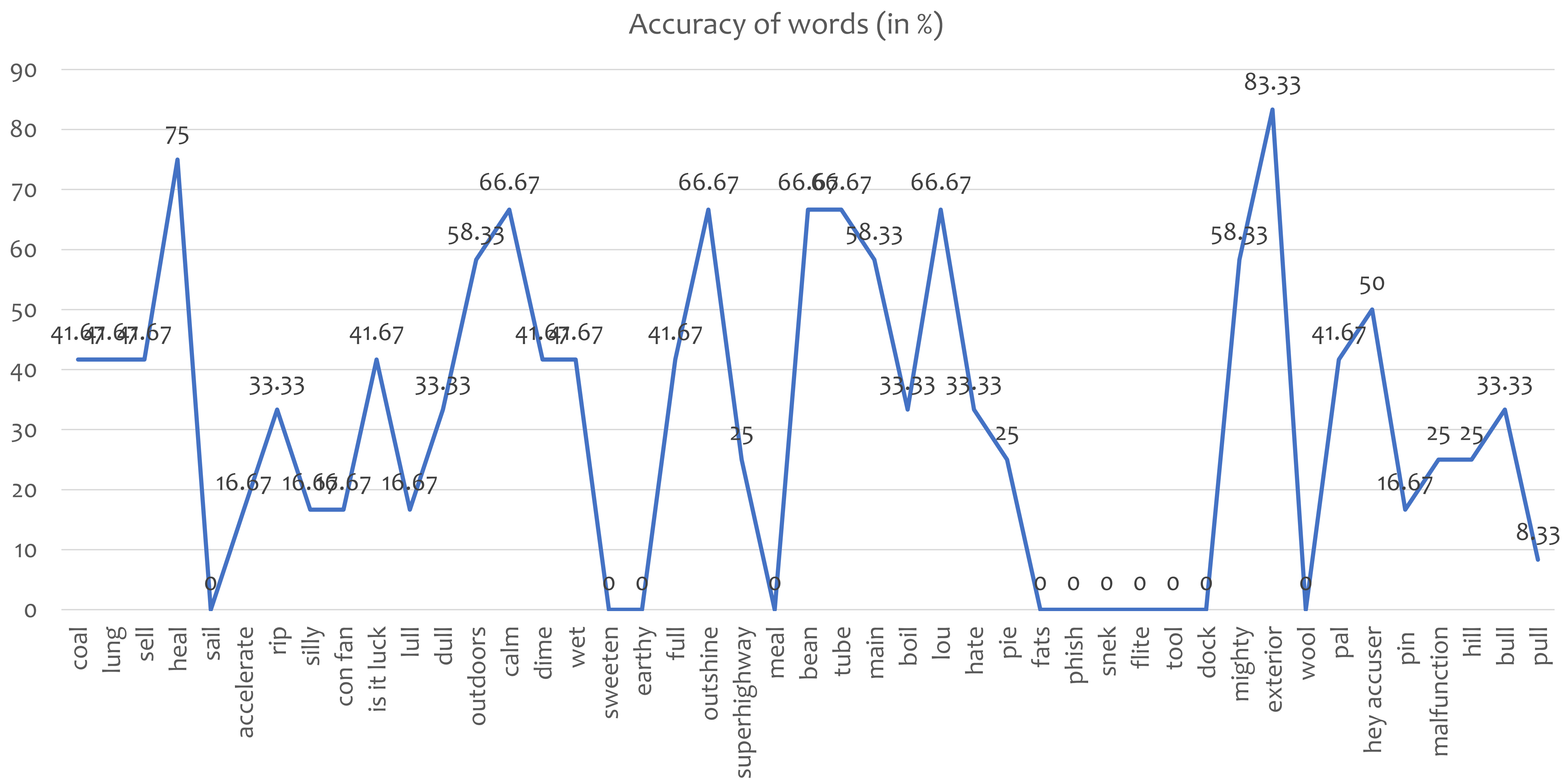
Intended Skill	Squatted Skill	Accuracy (%)	MCE
Coal Log	Call Log	41.667	Call
Lung Facts	Lang Facts	41.667	
Sell Size	Cell Size	41.667	
Heal Track	He’ll Track	75	Hell
Sail Across	Sale Across	0	
Accelerate Effect	Xcelerate Effect	16.667	
Rip Up	Rap Up	33.333	
Silly Mill	Silly No	16.667	
Con Fan	Khan Fan	16.667	Com
Is It Luck	Is It Lock	41.667	
Want to Lull	Want to Lol	16.667	
Get Dull	Get Doll	33.333	
Event Outdoors	Event Out Doors	58.333	
Calm Song	Com Song	66.667	
Dime Game	Time Game	41.667	

Predictable Errors

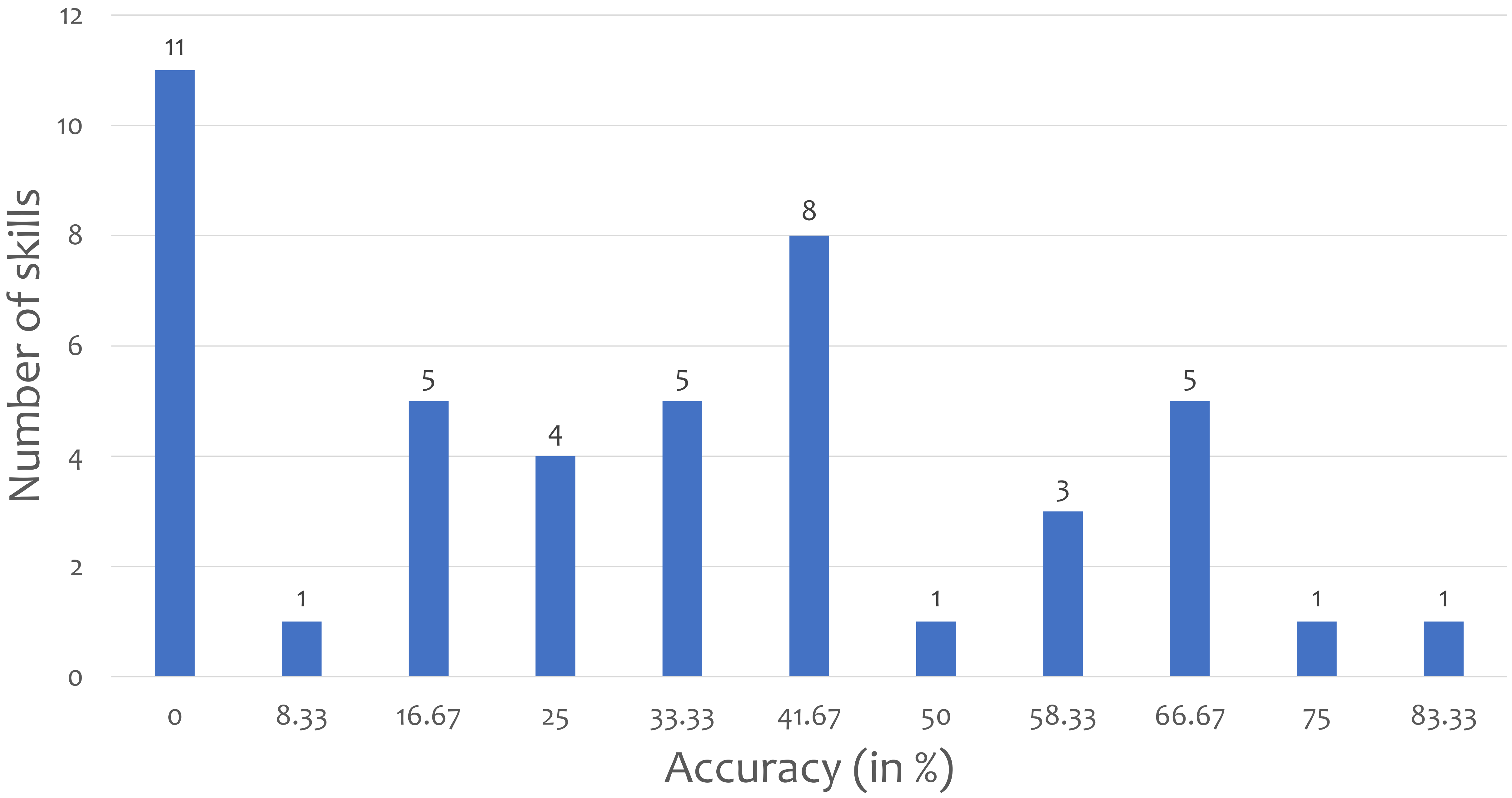
Intended Skill	Squatted Skill	Accuracy (%)	MCE
Wet Effect	What Effect	41.667	
Sweeten Facts	Sweden Facts	0	Sweet and
Earthy Lights	Fi Lights	0	
Full Season	Four Season	41.667	Fool
Outshine Band	Outshyne Bands	66.667	
Superhighway	Super Highway	25	
Cat Meal	Cat Meow	0	
Bean Bag	Been Bag	66.667	
Tube Clock	Two Clock	66.667	
Main Reasons	Maine Reasons	58.333	
Boil an Egg	Boyle an Egg	33.333	
Loud Noise	Louder Noise	66.667	
Hate time	Eight time	33.333	
Pie Fan	Pi Fan	25	
Animal Fats	Animal Facts	0	

Predictable Errors

Intended Skill	Squatted Skill	Accuracy (%)	
Phish Facts	Fish Facts	0	
Snek Memes	Snake Memes	0	
Flite Facts	Flight Facts	0	
Magic Tool	Magic Two	0	Door
Ship Dock	Ship Doc	0	
Mighty River	My T River	58.333	My TV
Exterior Angle	Xterior Angle	83.333	
Wool Facts	Well Facts	0	Bull
My Pal	My Pow	41.667	
Hey Accuser	Who’s There	50	
Pin Point	Pen Point	16.667	
The Malfunction	The No function	25	
Hill Rider	He’ll Rider	25	Hell
Bull Facts	Ball Facts	33.333	
Pull the Dog	Paul the Dog	8.3333	



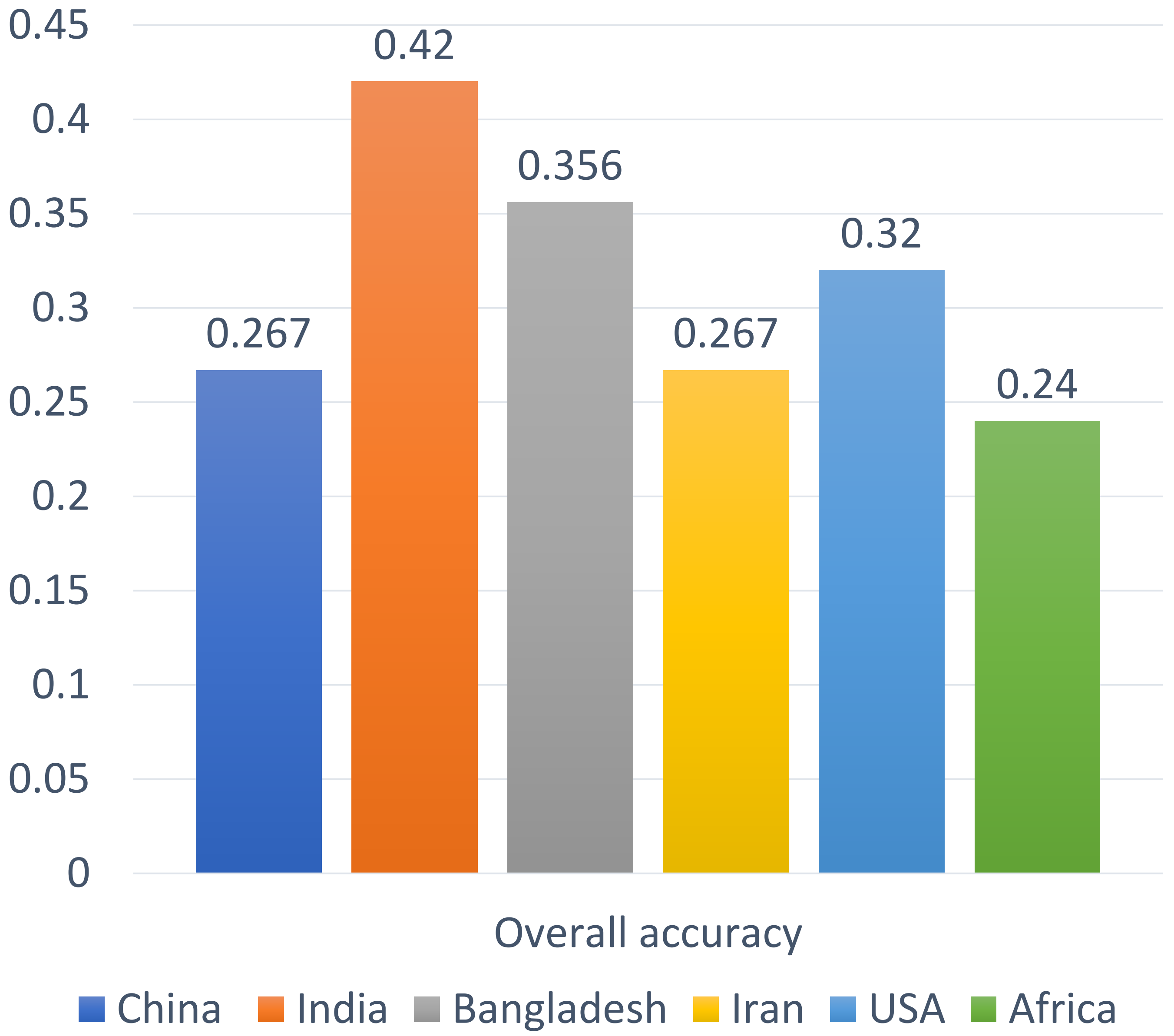
Accuracy of words (in %)



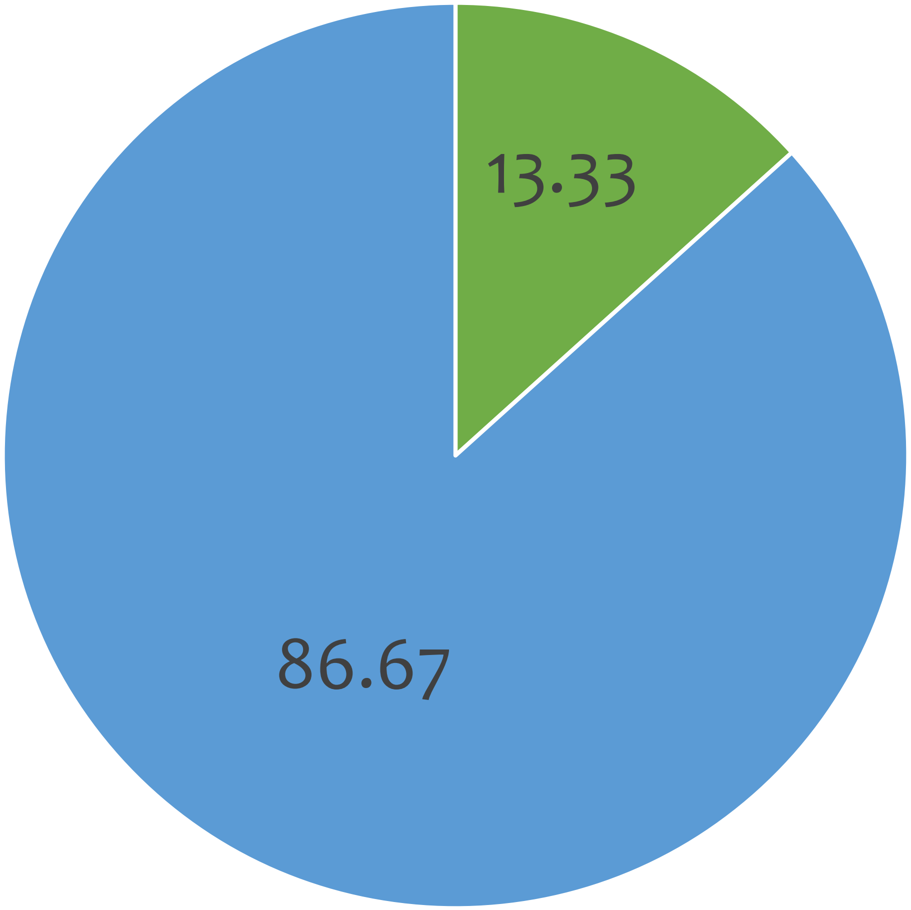
Number of skills
12
10
8
6
4
2
0
11
1
5
4
5
8
1
3
5
1
1
0
8.33
16.67
25
33.33
41.67
50
58.33
66.67
75
83.33
Accuracy (in %)

Do different regions exhibit unique predictable interpretation errors?

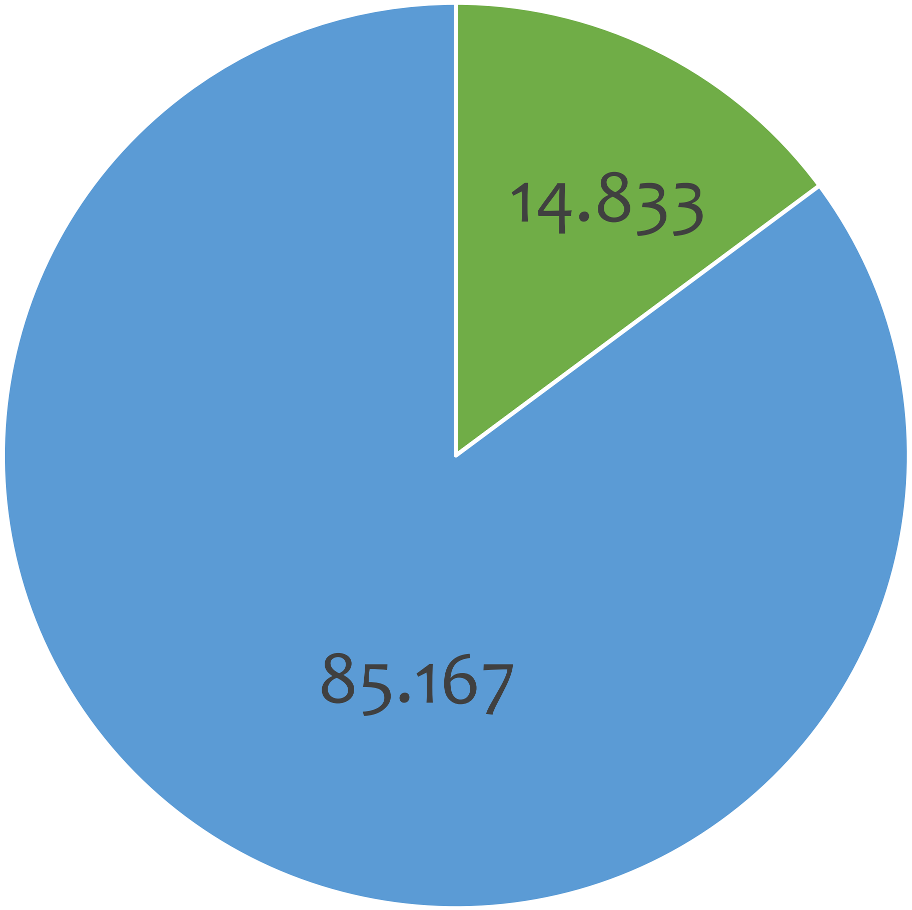
Regions


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.267
0.42
0.356
0.267
0.32
0.24
Overall accuracy
China
India
Bangladesh
Iran
USA
Africa

Females Vs Males - Accuracy in %

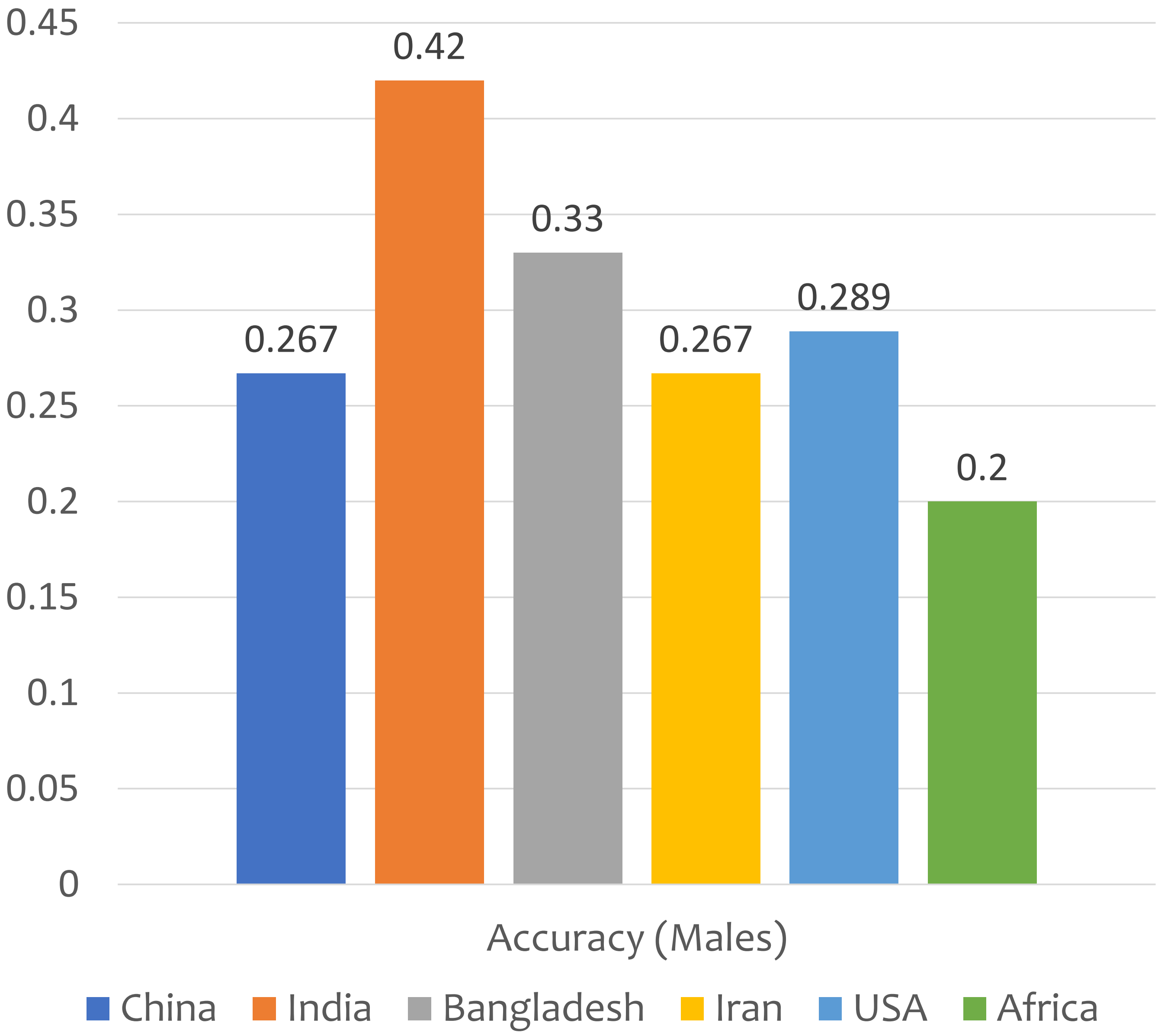


Males

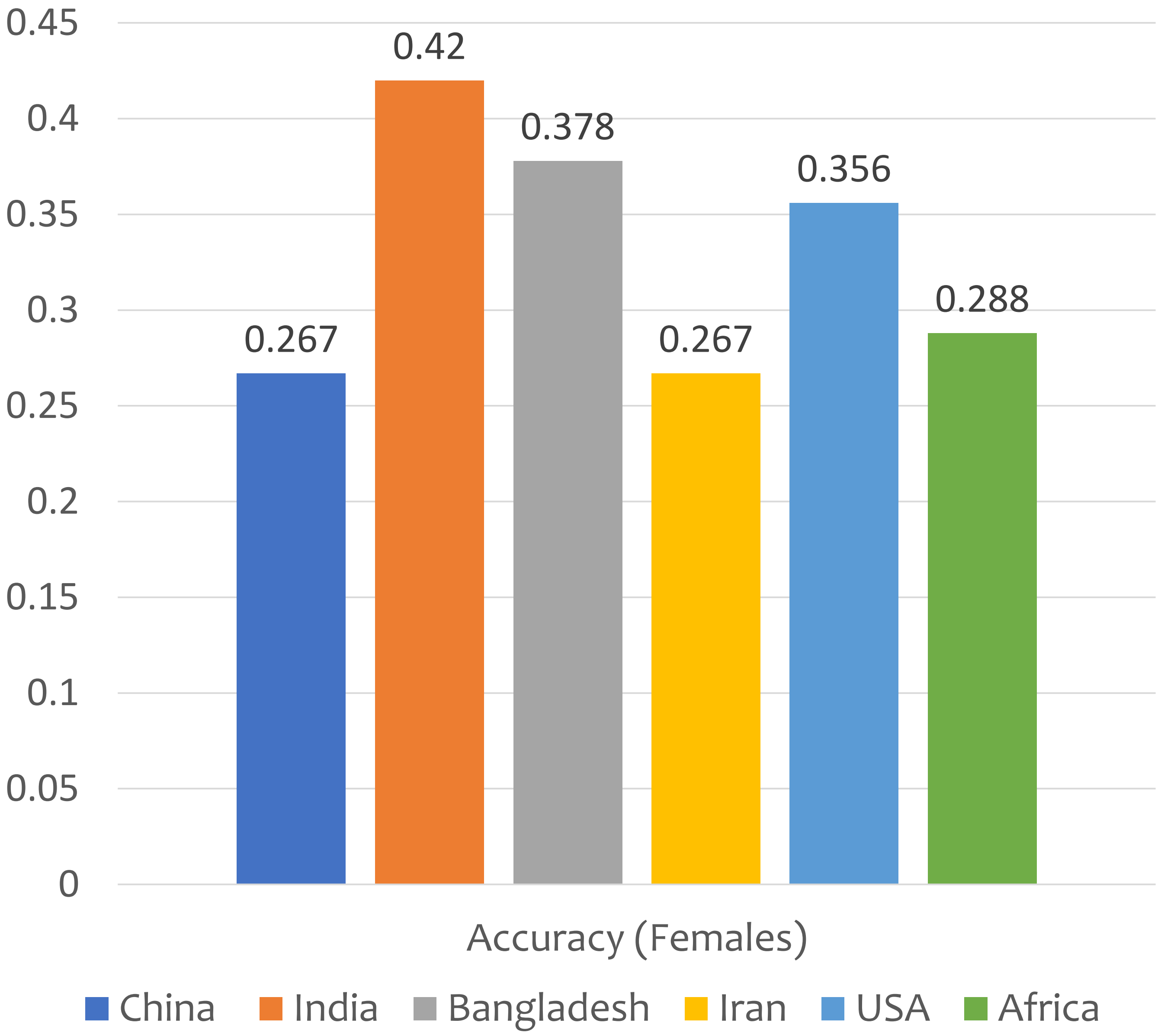


Females

Males


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.267
0.42
0.33
0.267
0.289
0.2
Accuracy (Males)
China
India
Bangladesh
Iran
USA
Africa

Females


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.267
0.42
0.378
0.267
0.356
0.288
Accuracy (Females)
China
India
Bangladesh
Iran
USA
Africa

Spear Skill Squatting Attacks

An attacker can leverage accent-specific predictable errors in Alexa to route distinct groups of users to skills that they didn't intend to go to

Alexa has different accuracies when interpreting speakers of different region

Alexa also has different accuracies when interpreting speakers of different genders

Results

Investigated speech samples

Some classes of interpretation errors are systematic

An attacker can leverage systematic errors to trigger malicious applications for users in the Alexa ecosystem

Ethical note

These skills are validated in a strictly developer environment. No skill was triggered and validated outside this environment.

No Real skills were targeted or tested.

Limitations

A core limitation of this analysis is the scope and scale of the dataset used in the analysis.

Another limitation of this work is that this is based on the key assumption that triggering skills in a development environment works similarly to triggering publicly available skills.

Alexa is a black box!

Future Scope

- Collecting richer dataset.
- Measuring harms of the skill squatting
- Investigation IoT trust relationships
- Generalizing our models

Thank You!