# CPSC 8570 Final Project Report - Skill Squatting Attacks on Amazon Alexa - A Demonstration

Nandini Krupa Krishnamurthy
*Clemson University*
*nandink@clemson.edu*

***Abstract*—Internet of Things (IoT) devices are already ruling the market. One of best examples of IoT are the smart speakers. The universe of smart speakers is dominated by Amazon Echo and Google home. Researches now say that by 2023 there would be 275 million voice assistant devices. In literal terms, they are going to exponentially grow by 1000%. So ideally we won't be controlling them , rather they would be controlling us. But these devices are of primary importance, not because they are so much developed in technology, but because they forego traditional input devices like keyboard and mouse and rely completely on voice input. Human voice becomes their primary input. There is a lot of effort put on these in the industry currently. Though these devices are improving in technology and making life easier, there are many cases where the users frequently face misinterpretations in everyday usage. There are many times, these VPAs perform actions other than what is intended to be performed. This error rate is even higher if the users are found to have an accent while speaking. What is more interesting is that there is a lot of difference in the way these VPAs interpret male and female voice messages. Keeping this as the basis, a group of researchers from University of Illinois Urbana-Champaign led a research to understand the basis of errors caused by these VPAs. Their work conducts an empirical analysis of interpretation errors in speech-recognition systems and investigate their security implications focusing on Amazon Alexa as the primary device. Their main concentration is on Amazon Alexa as this powers 70% of the smart speaker market.**

**In this project, a demonstration has been carried for alexa errors based on speech called as "Skill Squatting Attacks" [1]. These have been extended to the errors based on dialect and gender called as "Spear Skill Squatting Attacks". This also talks about the user study carried out for this project and also on the understanding of alexa errors classification. The project ends with a brief note on limitation and future work.**

## I. INTRODUCTION

Amazon wants to take over the world. In order to do this they built a platform where the third-party device developers can interface with Alexa directly. These are called as skills. So at the time we wrote the paper there were about 25000 skills on the Alexa skill store and this number was growing rapidly day by day [2]. There is a skill for almost everything. If a user wants facts about cats, there is a skill for that. If another user wants a lyft to go to the airport, there is a skill for that too. If anyone wants to know what is happening around the world, yes there is a skill for that. And you want to wake up early, and you are too lazy to set an alarm on your phone, then ask Alexa! In short, VPAs have made life very, very easy. Skills are to Alexa as Apps as to Android. The main difference lies in the usage of these apps. For Android apps, one needs to go to the App Store, download the app and then install it. There is involvement of the traditional input devices like a keyboard or Mouse. But, Alexa skills are enabled by default. This, in literal sense means that, if a user wants to get some information from Alexa or he just has to talk to Alexa and Alexa determines what skill is he trying to talk to and invokes it all by itself solely based on the voice input. This is a huge security implication because we don't know which skill is being invoked and Alexa is a complete Black Box in this example.

This gives rise to the topic of Skill squatting. Skill squatting can be defined as a type of attack when a malicious skill is invoked when the user calls for an action from Alexa. These malicious skills are unintended skills that are triggered without the user being aware of it. This is associated with the fact that Alexa is a black box. The core idea is straight forward. If an attacker knows what predictable errors Alexa is going to make when routing somebody to a particular skill, then all that he has to do is to build a skill that sounds like the intended skill. So, when a user tries to trigger a skill which was completely intended, they are instead routed to an unintended skill that is controlled by an attacker. This has some serious implications while doing bank transactions. Not only that, Alexa also has different interpretations for people of different regions and different genders. This is even more evident if the user has an accent. Therefore, 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.

Going on the same grounds, this class project for "CPSC-8570,Security in Advanced Networking technologies" aims to demonstrate "Skill Squatting Attacks on Amazon Alexa". It is also extended to demonstrate "Spear Skill Squatting Attacks" i.e., the attacks on Alexa based on ethnicity and gender, and this project also focuses on the security impacts of these two attacks. The main methods used, the case study that is done and also the results are elaborated in the following paragraphs. The Alexa skills are built on Alexa development environment and Lambda services are used which are again elaborated in the up-coming paragraphs.

## II. RELATED WORK

Skill squatting attacks on Amazon Alexa is a relatively new discovery. This work is a combination of various divisions - linguistics, human aspects of security and audio attacks on

VPAs. This project bases its work on the semantic misinterpretation of voice speeches. But this is not the only notable work in this field.

1) **Dialects in Speech**
Models on English speech is an old concept probably started around the 1970s. [3] There is also a recent phoneme and vowel data development to study the speech patterns depending on region and gender. [4] [5] [6]. Clopper worked on "semantic predictability" [7].
2) **Typosquatting and Human Factors**
As said above, this work aligns with human aspects of security. There is a long history for domain typosquatting [8] [9] [10] [11]. Notable works are that of Nikiforakis et al. who worked on finding vulnerable domain names based on homophone confusion [12], the investigation of Tahir et al. on some URLs being more prone to typosquatting than others [13].
3) **Other Skill Squatting Attacks**
There are variants to this work. Zhang et al. observed that Alexa favors the longest matching skill name when processing voice commands [14].
4) **Audio Attacks**
These types are one of the most common attacks on computing devices. A drone can malfunction and crash when an audio is given to it [15]. These have been used to bias sensor input on Fit-bit devices [16]. These attacks can also be used to steal private key information during key generation [17]. They can also be used to leak private data by modifying vibration sensors [18].

## III. Dataset

The original work used NSP (Nationwide Speech Project) Dataset which has the speech sample from six different dialect regions of the United States - Mid-Atlantic, Midland, New England, North, South, and West as shown in Figure 1.

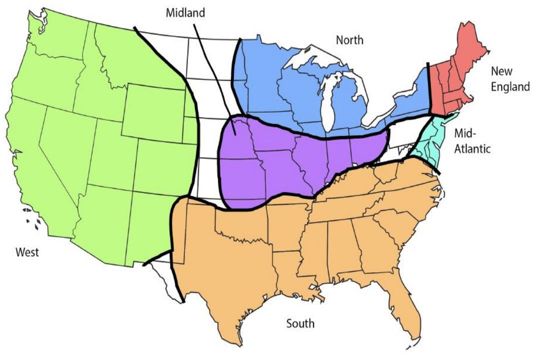
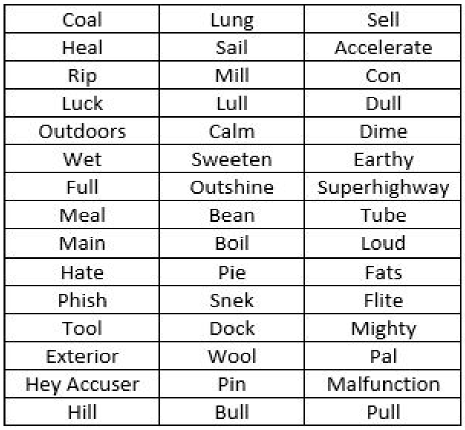


Fig. 1. Dialect-Regions in the U.S. [1]

NSP dataset contains of five male and five female speakers from these six geographical regions providing a set of 188 single word recordings. The errors from Alexa were found using these words. However, since the dataset was not available for the current project, the errors made by Alexa were used as the dataset for speech for a detailed study of errors. The paper provides a set of 24 words which are constantly mistaken by Alexa. This, along with the words from other papers [19] [20], a dataset of 45 words was created for demonstration purposes. These words are shown in the Figure 2.



| | | |
|---|---|---|
| Coal | Lung | Sell |
| Heal | Sail | Accelerate |
| Rip | Mill | Con |
| Luck | Lull | Dull |
| Outdoors | Calm | Dime |
| Wet | Sweeten | Earthy |
| Full | Outshine | Superhighway |
| Meal | Bean | Tube |
| Main | Boil | Loud |
| Hate | Pie | Fats |
| Phish | Snek | Flite |
| Tool | Dock | Mighty |
| Exterior | Wool | Pal |
| Hey Accuser | Pin | Malfunction |
| Hill | Bull | Pull |

Fig. 2. 45 words in our dataset

One female and one male from six different parts of the world – Africa, Bangladesh, China, India, Iran, USA – were asked to call out these words in the form of skills and the results were noted. This was carried out as a case study just to understand spear skill squatting attack better. So that became a dataset of 540 words. We will talk about the errors on these words and their interpretations by Alexa and the result in the forthcoming topics.

## IV. Architecture

Now let us look at Alexa architecture. The Figure 3 gives a simplified architecture of Alexa working.

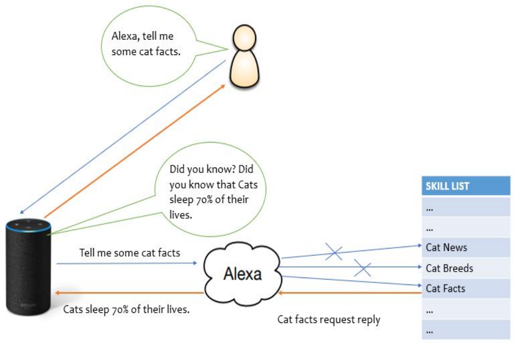

Alexa, tell me some cat facts.
Did you know? Did you know that Cats sleep 70% of their lives.
Tell me some cat facts
Alexa
Cats sleep 70% of their lives.
Cat facts request reply
SKILL LIST
...
...
Cat News
Cat Breeds
Cat Facts
...
...

Fig. 3. Alexa architecture

Let us see how Alexa works. Suppose the user wants some cat facts. He says, "Alexa, tell me some cat facts". Here

"Alexa" is the wake word. Wake word is a particular type of utterance that Alexa is looking forward to and on the utterance of which it starts listening and responding to the user. Amazon uses "Alexa", "Echo" and "Computer" as wake words for its devices. "Cat Facts" is the name of the skill called the invocation name, that the user intends to trigger and get some work done by Alexa. "Tell me" is the intent.

Now, the user asks Alexa if he can get some facts from "Cat Facts" Skills. Alexa transcribes it as "Tell me some Cat Facts" and send it to the Alexa Cloud. In this Cloud, let's say there are three skills that are related to Cats - "Cat News", "Cat Breeds" and "Cat Facts". Now Alexa looks for "Cat News", but this is not what is invoked, moves ahead to "Cat Breeds" but this is also not the intended skill. When it finds the "Cat Facts" skills, it routes the user request to these skills and seeks a response from this skill. "Cat Facts" Skill gives a response saying, "Cats sleep 70% of their lives". This response is sent to Alexa, which in turn reads it aloud for the user.

This is the ideal situation when everything works perfectly. Now let us assume that there is a skill squatting attack. Let us see what happens during this situation. Consider Figure 4. It talks about the architecture of Alexa system during skill squatting. The user says "Alexa, tell me some cat facts". But now, Alexa interprets it wrongs. It transcribes the speech as "Tell me some Cat Fax". Unfortunately, suppose say, there is a skill in Alexa Cloud as "Cat Fats" which is a malicious skill. Due to this inherent nature of Alexa, a wrong skill gets invoked as shown in the Figure 4. And Alexa responds saying something very irrelevant such as "Cat are not cats, they are dogs."

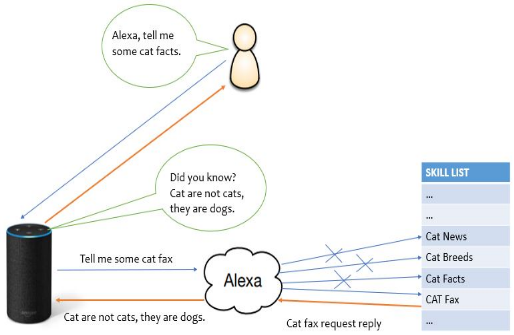

Alexa, tell me some cat facts.
Did you know? Cat are not cats, they are dogs.
Tell me some cat fax
Alexa
Cat are not cats, they are dogs.
Cat fax request reply
SKILL LIST
...
...
Cat News
Cat Breeds
Cat Facts
CAT Fax
...

Fig. 4. Architecture of Alexa when there is a skill squatting attack

But again, this is just a simple example. Skill squatting can be more serious during misinterpretations of names of skills related to Bank transactions. Consider a situation where an attacker knows that Alexa prefers character representations of words rather than the full representation of the word. So he builds a Skill called "Am X" in return to the "American Express" Skill "Amex". And this skill would definitely have all the features similar to the original skill. In this situation, there are chances that a user transaction may be attacked. This very similar situation is represented in the following Figure 5.

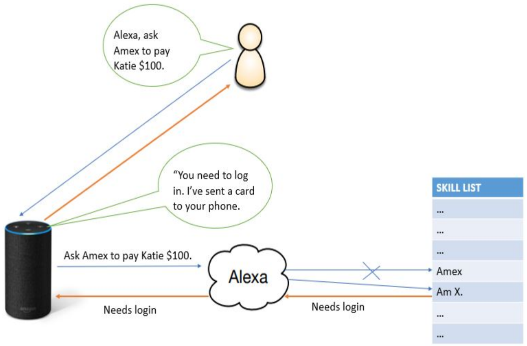

Alexa, ask Amex to pay Katie $100.
"You need to log in. I've sent a card to your phone.
Ask Amex to pay Katie $100.
Alexa
Needs login
Needs login
SKILL LIST
...
...
...
Amex
Am X.
...
...

Fig. 5. Demonstration of a probable attack on Amex - American Express Skill

## V. Methodology

### A. Assumption

The whole project has been implemented based on the following assumptions

1) Alexa is a black box
We do not know based on what does Alexa triggers a skill. We do not have an idea of what is the process on which Alexa decides which skill has to be triggered.
2) Similarity of work
This work relies on the assumption that triggering skills in a development environment is similar to that of triggering of publicly available skills.

### B. Ethical Consideration

1) Similar to the original work, this project uses skills that are validated in a strictly developer environment. No skill was triggered and validated outside this environment.
2) In the whole of the process no Real skills were targeted or tested.

### C. Usage

This section focuses on some of the common terms used in this project.

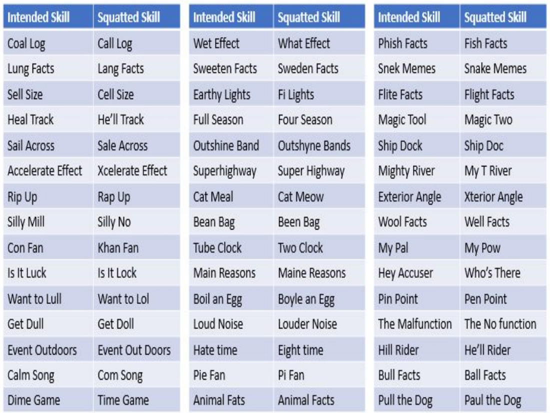
1) Actual/ Target/ Intended Skill
This the skill that was supposed to be invoked when the user called for it. When this skill is invoked, it is considered in this project that there is no attack over Alexa.
2) Squatted/ Unintended/ Malicious Skill
This is the skill that was invoked by Alexa due to misinterpretation of the audio speech. This was not supposed to be invoked and tells us that there can be probability that an attacker may utilize this error by Alexa to cause an attack over it to obtain some confidential data or to cause any harm.

3) MCE (Most Common Error)
When a misinterpreted word is invoked most of the times in place for a target word, that misinterpreted word is considered to be the MCE for the target word.

### *D. Process*

The whole process was divided into two phases – Building skills and Querying Alexa. Each of these stages have been elaborated below.

1) **Stage 1 - Building Skills**
The most important and the most time-consuming part was building skills. The name of the skills included a word from the dataset and another associated word. In simple term, Name_of_the_Skill = Word_from_the_Dataset + an_Assosiated_Word. For example, consider the first word from the dataset – coal. So a skill having the name “coal” in it was developed and was called as “Coal Log”. This was so done to make usage easier. Similarly, there are “Lung Facts” and “Sell Size” for second and third word in the dataset respectively. So totally 45 skills were built for all the words in the dataset. It was also necessary to build another set of 45 skills as squatted skills for the target skills. In simple words, each target skill has one squatted skill. For example, the target skill “Coal Log” has a squatted skill “Call Log”. On similar grounds, “Lung Facts” and “Sell Size” had the squatted skills “Lang Facts” and “Cell Size” respectively. A complete list of target skills and squatted list is shown in the Figure 6 below.

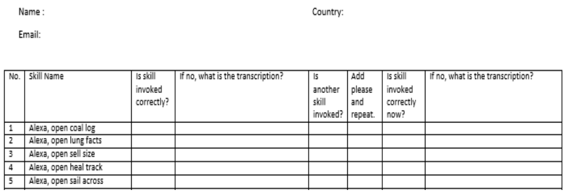


| Intended Skill | Squatted Skill | Intended Skill | Squatted Skill | Intended Skill | Squatted Skill |
|---|---|---|---|---|---|
| Coal Log | Call Log | Wet Effect | What Effect | Phish Facts | Fish Facts |
| Lung Facts | Lang Facts | Sweeten Facts | Sweden Facts | Snek Memes | Snake Memes |
| Sell Size | Cell Size | Earthy Lights | Fi Lights | Flite Facts | Flight Facts |
| Heal Track | He’ll Track | Full Season | Four Season | Magic Tool | Magic Two |
| Sail Across | Sale Across | Outshine Band | Outshyne Bands | Ship Dock | Ship Doc |
| Accelerate Effect | Xcelerate Effect | Superhighway | Super Highway | Mighty River | My T River |
| Rip Up | Rap Up | Cat Meal | Cat Meow | Exterior Angle | Xterior Angle |
| Silly Mill | Silly No | Bean Bag | Been Bag | Wool Facts | Well Facts |
| Con Fan | Khan Fan | Tube Clock | Two Clock | My Pal | My Pow |
| Is It Luck | Is It Lock | Main Reasons | Maine Reasons | Hey Accuser | Who’s There |
| Want to Lull | Want to Lol | Boil an Egg | Boyle an Egg | Pin Point | Pen Point |
| Get Dull | Get Doll | Loud Noise | Louder Noise | The Malfunction | The No function |
| Event Outdoors | Event Out Doors | Hate time | Eight time | Hill Rider | He’ll Rider |
| Calm Song | Com Song | Pie Fan | Pi Fan | Bull Facts | Ball Facts |
| Dime Game | Time Game | Animal Fats | Animal Facts | Pull the Dog | Paul the Dog |

Fig. 6. Target Skill and Squatted SKill

It is important for us to note how is the communication within Alexa skill. An Alexa skill consists of two main parts - Skill Interface and Skill Service. The skill interface records and transcribes the user input and maps them to the intents within the model. Intents are the predefined utterances. For example, in our architecture example, “tell me” was an intent. If an intent is detected, this interface created a JSON coded event and passes it to the Skill Service. The Skill Service determines what action has to be performed when it receives a JSON coded event. After this decision, it returns a JSON coded response to the interface. This speech response is sent back to the user through Echo.
Similarly, for the skills developed for this project, for simplicity, the intent used was “Open”. This was common for all skills. On triggering a Target Skill, it gave one of the following responses – “This is the Actual Skill”, “This is the Correct Skill”, “This is the Target Skill” or “This is the Proper Skill”. On triggering a malicious/unintended skill, one of the following responses are obtained – “Sorry, This is not the skill you asked for. This is an example of skill squatting attack. You should search for   instead”, “Sorry, This is not the Correct Skill. This is an example of skill squatting attack. You should search for ‘ ’ instead.” or “Sorry, This is not the Target Skill. This is an example of skill squatting attack. You should search for ‘ ’ instead.” The blank between “ ” is replaced by the name of the actual skill for that respective squatted skill. An important point to note here is that, these skills are not developed to give the user a result as described by its name. For example, the skill “Sweden Facts” does not necessarily give facts on Sweden. Rather, they try to educate the user if Alexa has interpreted correctly the word he/she has called out.

2) **Stage 2 - Querying Alexa**
Alexa has to be queried with all the names of the skills listed in Figure 6. Therefore a User survey was decided to be performed. A user survey form was created. The Figure 7 below shows the same.



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 | | | | | | |

Fig. 7. User survey form used for the experiment

Figure 7 only shows the first 5 of the 45 skills that the users were asked to invoke. For this a survey was conducted on the people from 6 different countries – Africa, Bangladesh, China, India, Iran and USA. There was a male and a female participant from each of the countries participating in this survey. Each participant had to invoke the skills in the order as given above and record the results for all the columns. The columns included – “Is skill invoked correctly?”, “If no, then what is the transcription”, “Is another skill invoked?”, “Add please and repeat”, “Is skill invoked correctly” and “If no, what is the transcription”. The last 3 columns were neglected due to time constraints of users for calling 90 skills together.

## VI. Observation and Results

This experiment upheld various interesting observations and surprising results. These observations summaries those on skill squatting and also spear skill squatting attacks.This section talks in detail about these observations.

1) Automatic enabling of Alexa skills may make Alexa prone to attacks.
   Since Alexa skills can be enabled by just by calling out their names, there is no security regarding how these skills get used. We do not even get to know which skill gets enabled or if the enabled skill is the target skill that we were looking for unless the user checks his Amazon account every time he triggers a skill which is cumbersome and practically infeasible.
2) Accuracy of these words is very poor.
   By accuracy we mean the number of times the Actual Skill was invoked and not the Squatted Skill.
   The average accuracy of the words ranged from 0% to 83.33% . Some words such as sail, sweeten, earthy, meal, fats, phish, snek, flite, tool, and dock were always misinterpreted. By misinterpretation it is not meant that a squatted skill that was created for the project was invoked, but the transcription of the word was always wrong. This gives rise to the thought that there are chances of an attack by usage of other unintended skills going beyond the concept of MCE. Figure 8 shows the accuracy of words.

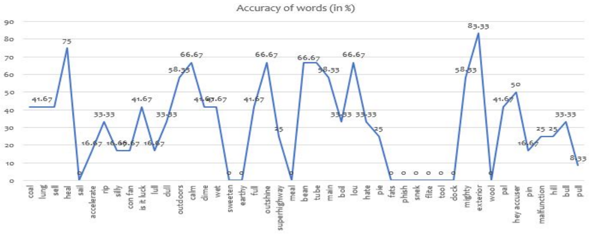


Fig. 8. Accuracy of words obtained from this experiment

3) Highest vs lowest accuracy
   There were 11 words with 0% accuracy (lowest) and only one word with 83.33% accuracy (highest). This means around 25% of the words in the dataset can easily be squatted almost all the time. Figure 9 shows number of words with accuracy range.
4) MCE
   The supposed to be squatted skills were built on the MCE of the target skills. But the experiment gave birth to new MCEs as listed below in the Figure 10. Note that MCE here only refers to the word from the dataset. MCE is the error of the word and doesn't denote the error of the skill.
5) Spear Skill squatting attack is evident.
   - It is in fact true that Alexa interprets people from different regions differently. This is particularly the case when a user has an accent. The Figure 11 shows the average accuracy of the Alexa interpretations by region. The highest accuracy was found to me by users from India and lowest was that of users from Africa. An important point to be noted here that, these results may not be accurate because users may get self-conscious while conducting the experiment. These results just show the accuracy of speakers among the people that participated in this user survey.

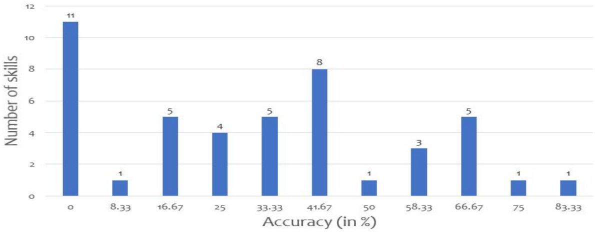
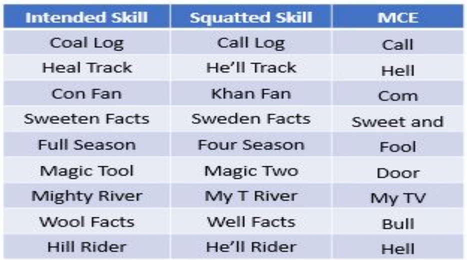


Fig. 9. Number of words with accuracy range



| Intended Skill | Squatted Skill | MCE |
|---|---|---|
| Coal Log | Call Log | Call |
| Heal Track | He'll Track | Hell |
| Con Fan | Khan Fan | Com |
| Sweeten Facts | Sweden Facts | Sweet and |
| Full Season | Four Season | Fool |
| Magic Tool | Magic Two | Door |
| Mighty River | My T River | My TV |
| Wool Facts | Well Facts | Bull |
| Hill Rider | He'll Rider | Hell |

Fig. 10. MCEs detected for the words

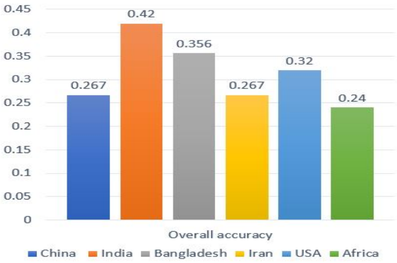


Fig. 11. Average accuracy of Alexa by region

- Females are indeed better interpreted than men by Alexa. Figure 12 shows these interpretations results by gender and by region as well.

## VII. Understanding Errors

The average accuracy of Alexa is around 31% which is very, very low. To understand better on how these errors might have raised, let us know the basis for these errors. These errors can

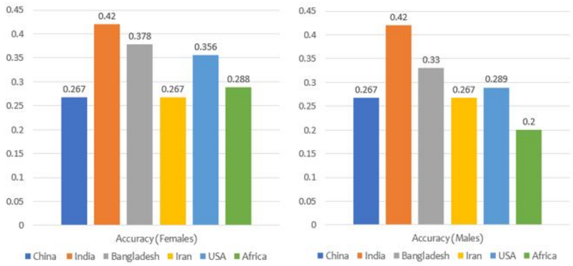


Fig. 12. Average accuracy of Alexa by region

be classified into four main groups - Homophones, Compound Words, Phonetic Confusion and Other errors [1]. These are described as follows.

1) **Homophones**
Homophones are those words which have the same pronunciation. In this project there were 11 words or approximately 24.4% of errors due to homophones. Some example are - "main" to "Maine", "boil" to "Boyle", "outshine" to "Outshyne", "Sell" to "cell", "bean" to "been", "pie" to "pi", "phish" to "fish", "exterior" to "xteriror", "accelerate" to "xcelerate", "flight" to "flite" and "dock" to "doc".
2) **Compound Words**
Alexa tends to break down words rather than returning a continuous compound word. In this project, there were 3 words or approximately 6.66% or errors due to compound words. For example, "superhighway" is split into "super highway" , "outdoors" is split into "out doors" and "mighty" is split into "my T".
3) **Phonetic Confusion**
A large number of errors attribute to the facts that there exist phonetic confusions between words due to the underlying phonetic structure. Here, 23 of 45 words i.e., approximately 51.1% were being misinterpreted due to phonetic confusions. There are "coal" to "call", "lung" to "lang", "heal" to "hell", "sail" to "sale", "rip" to "rap", "luck" to "lock", "lull" to "lol", "dull" to "doll", "calm" to "com", "dime" to "time", "wet" to "what", "sweeten" to "Sweden", "tube" to "two", "hate" to "eight", "fats" to "facts", "snek" to "snake", "tool" to "two", "wool" to "well", "pin" to "pen", "hill" to "hell", "bull" to "ball", "Con" to "Khan", and "pull" to "Paul".
4) **Other Errors**
There were some errors which could not be explained as to what was going wrong during interpretation of these. There were 8 such words or approximately 17.5% of errors could not be explained. These were "mill" to "no", "earthy" to "fi", "full" to "four", "meal" to "meow", "loud" to "louder", "pal" to "pow", "hey accuser" to "who's there" and "malfunction" to "no function".

## VIII. Limitations

Though this simple experiment was able to demonstrate skill squatting attacks, there were a few limitations to this attempt. They are explained in this section.

1) A core limitation of this project is the scope and scale of the dataset used. The word dataset consisted of only 45 easily squattable words. But it is evident that there is much more to this topic then these words. Not only this, the user study had only 12 participants. Though represented different regions of the world, there were only one female and one male representing a region which is inadequate to determine the accuracy.
2) Since the participants knew beforehand what was happening in this project and why are they being asked to give a survey, there might have been chances of them being conscious while speaking to Alexa. This might have changed the accuracy.
3) Another limitation is the assumption that triggering skills in a development environment works similar to triggering publicly available skills.
4) The fourth one is the assumption that Alexa is a black box. We do not know what exactly is the process behind the working of Alexa, so we would not be able to determine its exact work methodology.

## IX. Future Work

The results indicate that there is a lot of scope of improvement in the way alexa detects voice speeches. There are also several challenges to this project.

1) One of the important ones is the collection of better and richer dataset. The better the dataset, the better would be the interpretation of results.
2) This project demonstrates the skill squatting attacks successfully but does not measure the extent of harm caused by them. Another important future work remains to quantify the harms of this attack.
3) This project only concentrates on Amazon Alexa. But there are many IoTs and VPAs on the similar grounds serving as a platform for these attacks. Therefore another important challenge is to generalize these models.

## X. Conclusion

In this short demonstration, the skill squatting attacks on Amazon Alexa have been successfully demonstrated. This along with spear skill squatting attacks makes the essence of this project. There was also a user survey carried out to see and test the results of spear skill squatting attacks. These were done to show "how an attacker can leverage systematic errors to surreptitiously trigger malicious applications for users in the Alexa ecosystem" [1]. A separate dataset was created for this project which included those words which were commonly mistaken by Alexa. I conclude this project with a hope that this demonstration has made alexa users be more aware of the hidden attacks and provide a ground for those who can build on the more security for these VPAs.

# References

[1] Kumar, Deepak, Riccardo Paccagnella, Paul Murley, Eric Hennenfent, Joshua Mason, Adam Bates and Michael Bailey. “Skill Squatting Attacks on Amazon Alexa.” USENIX Security Symposium (2018).

[2] Alexa skills store. https://www.alexaskillstore.com/

[3] D. Crystal. Prosodic systems and intonation in English. CUP Archive.

[4] C. G. Clopper and R. Smiljanic. Effects of gender and regional dialect on prosodic patterns in American English. In Journal of Phonetics.

[5] C. G. Clopper and R. Smiljanic. Regional variation in temporal organization in American English. In Journal of Phonetics.

[6] W. Labov, S. Ash, and C. Boberg. The atlas of North American English: Phonetics, phonology and sound change. 2005.

[7] C. G. Clopper. Effects of dialect variation on the semantic predictability benefit. In Language and Cognitive Processes.

[8] P. Agten, W. Joosen, F. Piessens, and N. Nikiforakis. Seven months’ worth of mistakes: A longitudinal study of typosquatting abuse. In 22nd Network and Distributed System Security Symposium (NDSS).

[9] T. Moore and B. Edelman. Measuring the perpetrators and funders of typosquatting. In 14th International Conference on Financial Cryptography and Data Security.

[10] J. Spaulding, S. Upadhyaya, and A. Mohaisen. The landscape of domain name typosquatting: Techniques and countermeasures. In 2016 11th International Conference on Availability, Reliability and Security (ARES).

[11] J. Szurdi, B. Kocso, G. Cseh, J. Spring, M. Felegyhazi, and C. Kanich. The long ”taile” of typosquatting domain names. In 23rd USENIX Security Symposium (USENIX).

[12] N. Nikiforakis, M. Balduzzi, L. Desmet, F. Piessens, and W. Joosen. Soundsquatting: Uncovering the use of homophones in domain squatting. In International Conference on Information Security, 2014.

[13] R. Tahir, A. Raza, F. Ahmad, J. Kazi, F. Zaffar, C. Kanich, and M. Caesar. It’s all in the name: Why some urls are more vulnerable to typosquatting. In 13th IEEE International Conference on Computer Communications (INFOCOM).

[14] N. Zhang, X. Mi, X. Feng, X. Wang, Y. Tian, and F. Qian. Understanding and mitigating the security risks of voice-controlled third-party skills on amazon alexa and google home. Preprint, arXiv:1805.01525 [cs.CR], 2018.

[15] Y. Son, H. Shin, D. Kim, Y. Park, J. Noh, K. Choi, J. Choi, and Y. Kim. Rocking drones with intentional sound noise on gyroscopic sensors. In 24th USENIX Security Symposium (USENIX).

[16] T. Trippel, O. Weisse, W. Xu, P. Honeyman, and K. Fu. Walnut: Waging doubt on the integrity of mems accelerometers with acoustic injection attacks. In 2nd IEEE European Symposium on Security and Privacy (Euro SP).

[17] D. Genkin, A. Shamir, and E. Tromer. Rsa key extraction via low-bandwidth acoustic cryptanalysis. In 34th International Cryptology Conference (CRYPTO).

[18] N. Roy. Vibraphone project webpage. http://synrg.csl.illinois.edu/vibraphone/ .

[19] Zhang, Nan, Xianghang Mi, Xuan Feng, XiaoFeng Wang, Yuan Tian and Feng Qian. “Dangerous Skills: Understanding and Mitigating Security Risks of Voice-Controlled Third-Party Functions on Virtual Personal Assistant Systems.” SP 2019 (2019).

[20] Zhang, Yangyong, Lei Xu, Abner Mendoza, Guangliang Yang, Phakpoom Chinprutthiwong and Guofei Gu. “Life after Speech Recognition : Fuzzing Semantic Misinterpretation for Voice Assistant Applications.” (2018).