IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 March 2024, accepted 11 March 2024, date of publication 18 March 2024, date of current version 22 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3379141

RESEARCH ARTICLE

A Closer Look at Access Control in Multi-User Voice Systems

HASSAN A. SHAFEI[1,2] AND CHIU C. TAN[1], (Member, IEEE)
[1]Department of Computer and Information Science, Temple University, Philadelphia, PA 19122, USA
[2]Department of Computer Science and Information Technology, Jazan University, Jazan 45142, Saudi Arabia

Corresponding author: Hassan A. Shafei (tuj27065@temple.edu)

ABSTRACT Voice-controlled systems have revolutionized user interactions, making technology more accessible and intuitive across various settings. In multi-user environments, such as households, voice assistants like Amazon Alexa are favored as they enable seamless interaction with devices and services. However, the convenience these systems offer comes with challenges, especially concerning privacy and security. In environments where multiple users interact with the same voice assistant, the need for sophisticated access control mechanisms becomes apparent to prevent unauthorized access to sensitive information. This study assesses the effectiveness of voice access control mechanisms within these multi-user contexts, shedding light on the inherent privacy risks associated with shared voice-controlled systems. First, the study demonstrates vulnerabilities in the current access control mechanisms concerning users' private data. Second, a framework for automated testing is developed to explore the access control weaknesses and determine whether the accessible data is of consequence, as not all information may be equally sensitive or vital to users. Third, two flaws within the access control mechanisms offered by the voice system are identified, highlighting the susceptibility of existing access controls to unauthorized access. Finally, the study reveals that operations on the system are protected, whereas other operations that are not protected still reveal user's private information. These findings underscore the need for enhanced privacy safeguards and improved access control systems in multi-user environments. Recommendations are offered to mitigate risks associated with unauthorized access, focusing on securing the user's private data on the voice assistant.

INDEX TERMS Smart speakers, virtual personal assistants, voice interface, smart home assistant, access control, private information, privacy, multi-user, shopping data.

I. INTRODUCTION

The proliferation of Voice User Interface (VUI) has transformed how we live and interact with our living spaces. Using VUI in smart homes are no longer confined to the realm of tech enthusiasts; they have gained widespread popularity, and the number of connected services continues to grow at an astonishing pace [1]. Smart home assistants known as smart speakers, such as Amazon Alexa and Google Home, are increasingly popular and entering tens of millions of households [2], [3], [4], [5], offering an array of services, including online shopping, managing to-do lists, playing music, controlling objects and devices around the houses, making calls, sending messages and many more [6], [7], [8], [9], [10], [11], [12].

A. MULTI-USER ENVIRONMENT CHALLENGES

Smart speakers powered by voice assistants (VAs), located in common spaces, serve multiple household members, encompassing family, friends, and occasional visitors. This shared usage model necessitates the ability of the VA apps to recognize and differentiate between users, ensuring personalized and secure interactions. This becomes particularly critical in the context of various voice-activated real-time applications, where user-specific preferences and data security are crucial. These voice real-time applications include:

- Reminder/calendar applications. In a multi-user environment, these applications could manage updates and

The associate editor coordinating the review of this manuscript and approving it for publication was Martin Reisslein.

notifications for multiple users based on individual preferences and permissions. The VA must distinguish between users, allowing them to view or edit only their appointments or reminders. At the same time, parents or household heads might need broader access to coordinate schedules for all family members. For example, while one family member may have access to add or modify all calendar entries, another may only view them.

- Retail applications. It involves the integration of VA and shopping services to enhance shopping experiences in online shopping. These applications allow users in a multi-user environment to purchase, search for products, and perform other retail-related tasks using the VA. The VA must distinguish between users to prevent unauthorized access to payment information or personal purchase history. For instance, one user might have permission to make purchases, while another can only add items to a shopping list or shopping cart.

While we have highlighted reminder/calendar and retail applications due to their impact on multi-user environments, it is essential to recognize that the potential for future applications is evolving. As the VA devices are not individual-like mobile phones but are placed in communal spaces, we anticipate the emergence of additional applications tailored to these environments in the future [13], [14], [15], [16]. These future applications will likely extend beyond individual use, requiring even more sophisticated user identification and access control mechanisms to ensure privacy, security, and personalization for all users in a household.

Real-time applications highlight the necessity for robust user identification and access control mechanisms. For instance, a shared VA device may grant access to sensitive information. The current access control implemented in VA devices is a coarse-grained (all or nothing) access control where all users have full or no access [17]. Amazon VA (Alexa), for example, might schedule events or read the primary user's texts. However, given the coarse-grained mechanism, anybody within the home environment, including children, guests, or roommates, can access this information.

VA applications in a multi-user environment should ideally support two criteria: (1) it must provide personalized experiences for different users with different expectations in maintaining privacy, and (2) have a balance between usability and security. Having such criteria is important in a multi-home environment. For example, app services facilitated by VAs, such as online purchasing, scheduling, record-keeping, and planning, allow all users in a household legitimate access to initiate various commands [18], [19], [20], [21]. The VAs would respond with information, some of which are sensitive. We want to ensure and protect user privacy. Privacy concerns are crucial in such settings; sensitive personal data can be exposed to unintended recipients. These issues underscore the need for robust user identification and access control mechanisms in VA systems to ensure that they operate securely and appropriately in a multi-user environment with multi-access.

A straightforward approach to safeguarding user privacy is requiring users to log in every time they use the service. For example, every time a user accesses a calendar app, the user would be required to enter a PIN to retrieve information. This method addresses problem (1); however, it is inconvenient because it disrupts the quick interaction that is one of the primary benefits of using VAs. Requiring a PIN for every action can significantly slow down the process, rendering the system less user-friendly and accessible, especially for tasks that may not necessitate high security [14], [22], [23], [24]. Alternatively, the use of voice authentication could seem to address both problems (1) and (2) as it provides a method for verifying user identity without manual input. However, this approach is error-prone and can struggle under real-life conditions such as background noise or varying speech patterns, making it unreliable and potentially burdensome [25], [26], [27].

The challenge between usability and security in a multi-user access environment is significant. On the one hand, there is a need for these systems to be highly usable, with minimal barriers to access, ensuring that users can interact with them seamlessly in their daily routines. For instance, in using a VA for financial inquiries, a user may need to promptly check whether a bill payment has been made. On the other hand, enhancing security measures, such as implementing stringent user authentication processes for every action a user performs, can potentially hinder the ease of use and immediate responsiveness that characterize smart speakers. For example, certain actions like controlling the light do not necessitate stringent security protection. The balancing act between convenient access and safeguarding user privacy and security is crucial. An overemphasis on security could lead to cumbersome user experiences, while prioritizing convenience might expose the system to vulnerabilities in exposing user information.

B. CONTRIBUTIONS

Considering the challenges identified with voice applications that cater to accommodating multiple users, providing personalized experiences with varying privacy expectations, and maintaining a balance between security and ease of use, the focus of this study was directed towards an application that meets these criteria. The online shopping app was selected for examination as it is widely popular and well developed [28], [29], [30], [31]. Online shoppers increasingly use voice assistants to make purchases. VAs have become a common choice for online shopping as they offer a convenient, hands-free experience [32], [33]. Several major companies, including Amazon, Walmart, Costco, and Target, are already leveraging VA to gain a competitive advantage by providing voice shopping services to customers [34], [35], [36], [37]. The potential for exposure of sensitive shopping data, such as personal products and medication data, is heightened. Moreover, the ease of placing orders or

accessing shopping functions with simple voice commands can lead to unauthorized purchases or privacy breaches [38], [39], [40], [41], [42].

Given the widespread adoption of smart speakers, this paper focuses on examining VUI from the perspective of Amazon smart speakers. The choice of Amazon Alexa as the primary platform is deliberate and stems from its status as one of the most extensively used voice interfaces globally [43]. Amazon Alexa's popularity ensures relevance to a substantial user base, making it a representative choice for investigating voice systems' security and privacy aspects. Specifically, a case study was conducted using the shopping service provided by Amazon through its Alexa assistant, exploring the potential exposure of a user's private shopping data in smart home environments. While this study is platform-specific, the insights gained can contribute to a broader understanding of security challenges in voice-controlled systems. The contributions are summarized as follows:

- This study shows that the current access control allows for the retrieval and addition of shopping data. This paper proposes a framework for automated testing to explore the shortcomings of the access control to assess whether the accessible data is of consequence, as not all information may be equally sensitive or vital to users. This study shows that critical operations like deleting products, making purchases, and updating the payment method are protected. In contrast, the other operations, such as retrieving and adding products that are not protected, still reveal user information.
- This work identifies two significant flaws within the access control mechanisms proposed by the VA. The first flaw allows for the complete exposure of in-cart and list shopping products and access to purchase history for periods shorter than a month. The second flaw facilitates a targeted exploitation based on specific product details, such as product title and brand, to reveal the user's purchase history.

The remainder of this document is structured as follows: Section II discusses related work in the field. Section III introduces the background on access control protection provided by voice systems, challenges in testing voice assistants, and the adversary model. Section IV presents the design of the testing framework. Section V details the results obtained from various interaction strategies. Section VI offers recommendations to mitigate the risks of unauthorized access. Finally, Section VII concludes the paper with key findings and contributions of this work.

II. RELATED WORK

Smart homes and voice assistants raise significant concerns regarding security and privacy vulnerabilities, as demonstrated by previous user studies that delved into user concerns and perceptions surrounding the smart home environment [13], [14], [18], [19], [19], [44], [45], [46], [47], [48]. While the literature survey outlines the landscape of prior research in smart home security and voice assistant privacy, this work distinguishes itself in several key aspects.

TABLE 1. Summary of literature survey (SH: smart home, vs: voice system, adv.: adversarial, A.C.: access control).

Prior Work	Domain	Multi-user Multi-device Environment	Adv. Audio Attack	A.C Adversary Model
Zeng et al. [18]	SH	✓	×	×
He et al. [45]	SH	✓	×	✓*
Yuan et al. [47]	VS	×	✓	×
Carlini et al. [59] [60]	VS	×	✓	×
Zhang et al. [61]	VS	×	✓	×
Our Work	VS	×	×	✓

* Access control focuses on IoT capabilities

Table 1 provides an overview of these works, categorizing them based on the domain, consideration of multi-user, multi-device environments, adversarial audio attack focus, and the inclusion of an access control adversary model.

A. SMART HOME SECURITY AND PRIVACY

In the field of smart home privacy and security, substantial prior research emphasizes limiting malicious activities, which focus on device access control and authentication for single-user scenarios [49], [50], [51], [52], [53], [54], [55], [56]. He et al. [45] give a detailed smart home user study that depicts users' concerns about fine-grained access control in multi-user smart environments. Zeng et al. [18] describe their findings about smart home security and privacy problems. Both studies express smart home users' concerns about the need for a smart home access control system. Furthermore, these studies synthesize many design standards for access control mechanisms that reflect users' needs. Matthews et al. [57] also raise concerns about smart home users who use the same devices and accounts. Garg et al. [58] investigated the restrictions of sharing smart home devices among users with diverse social relationships in a recent user survey and highlighted future design requirements for smart house access control. While prior works have explored user concerns about in-home privacy, they have largely overlooked access control mechanisms provided by service providers in multi-user environment voice systems. In contrast, this research involves conducting experiments that reveal weaknesses in the access control measure implemented by Amazon, shedding light on its inability to protect user's private data from unauthorized access in multi-user environments voice systems.

B. VOICE SYSTEM SECURITY AND PRIVACY

In the domain of voice assistant privacy and security, a major source of concern pertains to the potential for security and privacy violations associated with voice assistants. Several studies [14], [44], [46], [62], [63], [64], [65], [66], [67] were conducted to understand users concerns regarding voice assistant systems. A large body of existing works focuses on adversarial audio commands, particularly those transmitted remotely to manipulate voice assistant interfaces and execute malicious actions. For instance, Yuan et al. [47] introduced Commandersong, a technique embedding voice commands

within songs to render them imperceptible to human listeners. Similarly, [59] crafted malicious audio that mimics ordinary speech but is interpreted differently by targeted devices. Additionally, studies such as Hidden Voice Commands [60] and Dolphin Attacks [61] pioneered adversarial audio attacks designed to exploit vulnerabilities in speech recognition systems while remaining acoustically undetectable to human ears. These investigations have unveiled the potential threats posed by malicious actors seeking to compromise speech recognition systems and manipulate voice assistants for nefarious purposes. Although these studies shed light on potential threats, they seldom explore the dimension of in-person or remote can issue commands to gain unauthorized access in multi-user environments. This research fills this gap by investigating how such adversaries might exploit voice systems to access and retrieve users' private information.

C. MULTI-USER ACCESS CONTROL

A line of research on multi-user access control has explored and proposed different access control strategies when multi-users share a single or multiple devices. Liu et al. [68] proposed xShare, a user access framework for the mobile phone ecosystem that enforces policies on file-level accesses. Ni et al. [69] introduced DiffUser, an access privilege-based user access management paradigm for the Android environment, which is only practical to a single device. In a smart home environment, Zeng et al. [13] developed an access control prototype with several control options for users in the smart home. They evaluated four distinct access control techniques in a month-long user study of seven households to better understand the users' demands and enhance the design. Although they developed a proof-of-concept framework for conducting a user study and outlining the access control requirements in smart home environments, they did not implement it in real-world systems. They did not account for user voice interaction in a multi-user smart home environment. Schuster et al. [70] developed situation-based access control, considering several environmental parameters. The authors examined the device's condition and the users' location to determine a valid access request. Yahyazadeh et al. [71] proposed Expat, a policy language for defining policies based on user demands. However, this study diverges from these paths. The research is dedicated to scrutinizing the access control mechanisms currently deployed in real-world voice systems. This work aims to uncover the existing shortcomings and limitations within these systems, particularly focusing on multi-user environments. The investigations reveal that while theoretical and prototype models provide valuable insights, there remains a gap in their application to real-world scenarios, especially regarding voice interaction in multi-user settings. This study does not implement new access control policies; rather, it analyzes and highlight the vulnerabilities of current implementations, emphasizing the service's inability to safeguard user's private data from unauthorized access in shared environments.

III. BACKGROUND AND ADVERSARY MODEL

This section discusses the access control protection offered by Amazon Alexa to secure user shopping data, as discussed in III-A. The complexity of voice assistant systems and the inherent challenges in testing such systems are subsequently discussed in III-B. Then, this work assesses the effectiveness of Amazon Alexa's access control mechanisms in safeguarding user privacy and data, as outlined in III-C. Finally, the focus is on understanding potential methods an adversary might employ to access sensitive user information, detailed in III-D.

A. ACCESS CONTROL PROTECTION

Amazon, as a service provider, offers various access control protection measures to ensure the security and privacy of user shopping data.

1) EXTERNAL DEVICE VALIDATION

There are operations that require user action. Users need to perform certain operations, such as changing the delivery address and payment methods, using the Amazon app or website [72], [73]. In these cases, Alexa would send users a wild card through the Alexa app or instruct them to use the Amazon app/website to perform them. By mandating the users to use the app or website, Amazon aims to fortify the protective shield around user accounts and sensitive personal information. This approach serves as a robust safeguard against unauthorized access and potential breaches, assuring users that their data remains confidential.

a: CONFIGURATION SETUP

Users are required to use the Amazon app or website to complete specific operations, thereby contributing to the overall security of their accounts and personal information.

2) REQUIRED PASSWORD

To enhance the security of Amazon Alexa, the primary account holder can set up a 4-digit voice password for placing orders and making purchases. The primary account holder can establish a unique 4-digit voice password, ensuring that only individuals with knowledge of the voice code can place orders. This feature is designed to prevent unauthorized purchases. This added layer of security helps prevent unauthorized purchases and access to sensitive information [74].

a: CONFIGURATION SETUP

The configuration of the 4-digit voice password takes place within the Alexa app. The primary account holder is responsible for configuring a 4-digit voice password, restricting order placement to those with knowledge of the voice code.

3) VOICE PROFILE

Amazon Alexa offers some protection that allows the primary account holder to set up *only* recognized voice profiles

could place orders through the voice profile features. The Alexa voice profile is a feature within Amazon's voice assistant ecosystem that allows individual users to create a unique voice profile linked to their account. The primary purpose of Alexa voice profiles is to provide a more personalized and secure experience when interacting with Alexa-enabled devices. This feature offers an additional layer of security, helping to prevent unauthorized access to personal information and settings. Alexa can tailor responses, content, and recommendations to that specific user, including personalized news briefings, music playlists, and calendar events. In the context of Amazon purchases, the primary account holder has the authority to manage users with Voice IDs, allowing control over purchasing permissions. This ensures that only recognized voice profile members can place orders through Alexa. Amazon does not explicitly design the voice profile for security purposes like the password. However, it provide some access control, such as the primary account can configure what voice ID can place orders [75].

a: CONFIGURATION SETUP

The primary account holder can set up, manage and configure voice profiles within the Alexa app, controlling who can place orders through voice commands.

4) OTHER SECURITY CONSIDERATIONS

In addition to access control measures, users can change the wake words used to activate the device when using Amazon Alexa. Users have the option to change the wake word from the default "Alexa" to alternatives like "Amazon," "Echo,", "Computer" or "Ziggy" via the Alexa mobile app or website [76]. However, it is important to note that any person who knows the wake word assigned to an Alexa device can issue voice commands, potentially leading to unauthorized access if the wake word is discovered.

B. VOICE ASSISTANT TESTING CHALLENGES

The current voice assistants complicates the testing process since the system is hosted on the voice assistant cloud service as a black box, as depicted in Figure 1. Thus, If an error occurs during testing, the voice interface typically responds with a default response (*e.g., "I'm sorry, I didn't quite get that")* and exits. This error could be a result of background noise, the way the command was spoken, the user's accent, network problems, or the complexity of the underlying system components.

The common mechanism for interacting and accessing this array of services (e.g., queries, services, smart devices, shopping) in a household is through voice commands. A user would speak naturally, such as *"turn on the light in the kitchen."* However, the complexity of the system arises from the need to process these voice commands, which involves multiple components, including Automatic Speech Recognition (ASR), Natural Language Processing (NLP), and Natural Language Understanding (NLU). These

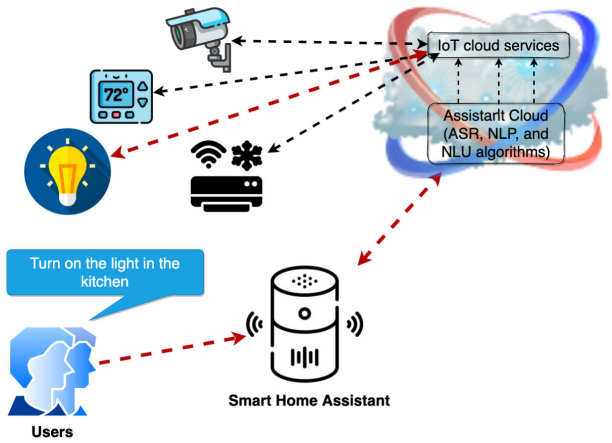

FIGURE 1. Voice interaction in modern households where users can use natural voice commands to control various devices and services.

components work together to process user requests on the smart home assistant cloud, extracting user intent and forwarding it to the corresponding IoT cloud service [8], [77], as shown in Figure 1. The VA act as intermediaries, translating voice commands into IoT/service instructions and eliminating the need for a dedicated IoT hub to perform voice processing algorithms. The nature of voice interaction makes it challenging to test.

C. HOW WELL DOES IT WORK?

This work conducted an experiment of the access control implemented by Amazon Alexa that would prevent unauthorized access. This work did not explore access control mechanisms that require users to employ external validation means, such as the Amazon app or website, or rely on a knowledge-based password. These mechanisms are considered out of scope for a voice-based-activated system. Additionally, this work did not perform experiments related to other security considerations, such as changing the wake words, since they are generally considered less secure, given that Alexa wake words are publicly available [44], [78], [79], [80], [81], [82]. The experiments were conducted on the Alexa **Voice Profile** mechanism and see whether there is an access control for the following operations: *retrieving*, *adding*, *purchasing*, *deleting* products/items from the shopping cart and list. The primary objective is to determine the effectiveness of Alexa's access control measures in preventing unauthorized access to the contents of the shopping cart and shopping list.

1) EXPERIMENTAL SETUP

Multiple distinct Amazon voice profiles were created, adult1 (male, primary account holder), adult2 (female), adult3 (male, different age), computer generated voice(robotic), and AI generated voice (sound realistic to human). Each voice profile was configured following Alexa's provided instructions. The ability of Alexa to correctly identify each user was assessed by prompting with the question, *"Alexa, who am I?"* To which, Alexa responds by stating, *"I'm talking*

TABLE 2. Alexa access control: Voice profile experiment on shopping cart and list. It shows the retrieve, add, delete and place order on the shopping cart. Note, the shopping list can be used only to retrieve, add, and delete. Users are not allowed to place order on the shopping list.

Voice Profile	Shopping Cart				Shopping List		
	Retrieve	Add	Delete	Place order	Retrieve	Add	Delete
Adult 1 (primary)	✓	✓	×	✓	✓	✓	✓
Adult 2	✓	✓	×	×	✓	✓	✓
Adult 3	✓	✓	×	×	✓	✓	✓
Comp. Voice	✓	✓	×	×	✓	✓	✓
AI Voice	✓	✓	×	×	✓	✓	✓

to [user name]." For the computer-generated voice (robotic), text-to-speech provided for Mac users, known as Spoken Content, was utilized [83]. For the AI-generated voice, the pyttsx3 library was used for text-to-speech conversion [84]. Once the voice profiles were correctly configured, the Alexa Voice ID account settings were adjusted for purchasing, allowing recognition solely for adult1 to make purchases and place orders, while recognition for the other voices (adult2, adult3, computer, and AI-generated voice) was disabled.

2) EXPERIMENTAL PROCEDURE

The users with voice profile were utilized to interact with the Alexa smart speaker to perform operations such as *retrieve*, *add*, *place*, and *delete* products. Prior to engaging with Alexa VA, the following command was used, *"Alexa, who am I?"* followed by issuing the commands. This step was undertaken to verify that the VA recognize the voice profile and whether it allowed the voice profile to issue the command.

The voice profile users (adult2, adult3, computer, and AI-generated voice) were instructed to *add* Amazon shopping products to the shopping cart and list. Subsequently, other operations were performed such as *retrieving* information (*Alexa, what is in my shopping cart/list*), For *deletion*, the following commands were employed, (*"Alexa, delete/remove/erase product [title] from my shopping cart"*) or (*"Alexa, delete/remove/erase product [title] from my shopping list."*) For *placing* orders and making purchases through shopping cart, commands such as (*"Alexa, place/confirm order" or "Alexa, checkout my Amazon cart.")* were used.

3) RESULTS AND OBSERVATIONS

The outcomes of the experimentation on Alexa's access control voice profiles, as illustrated in Table 2, provide valuable insights into the platform's security measures. For each voice profile, including Adult 2, Adult 3, Computer Voice, and AI Voice, Alexa allowed the *retrieval* and *addition* of items to both the shopping cart and shopping list. However, it is essential to highlight that none of the voice profiles were authorized to *delete* products from the shopping cart. The experiments reveal that Alexa's access control mechanisms do not effectively prevent unauthorized access to the contents of the shopping cart or shopping list. Specifically, operations such as *retrieval* and *addition* were permitted for users other than the primary account holder, indicating a gap in the control system. It is noteworthy that Adult 1 was acknowledged to place orders, indicating the effectiveness of access control for recognized users, as documented by Amazon. However, this recognition does not extend to preventing access to shopping data.

Takeaways: Alexa permits retrieval and addition of items to the shopping cart and shopping list for all voice profiles tested. However, none of the voice profiles, including the primary account holder, can delete products from the shopping cart. Deletion remains operational for the shopping list. These findings indicate that Amazon's access control measures are insufficient in preventing unauthorized access to shopping cart contents.

D. ADVERSARY MODEL

Previously, the shortcomings of the access control provided by Amazon Alexa were demonstrated. Next, this study will delve into the potential privacy implications resulting from these vulnerabilities. The aim is to determine whether the information that can be revealed due to these limitations is actually important and of consequence, as not all information is equally sensitive or vital to users.

The adversary objective is to gain unauthorized access to the user's private information, including confidential data such as the shopping content. To clarify, *Private information* refers to data that is considered confidential and specific to the user, typically not intended for public disclosure. This encompasses a wide range of items, including personal medication, products related to sex and pleasure, and other categories. What constitutes private information can vary from individual to individual. It is essential to acknowledge that privacy is a highly subjective concept.

The products falling into these categories have been described in prior research as "unmentionables," "socially sensitive products," or "controversial products" [38], [39], [40], [41], [85], [86]. These terms collectively characterize items that, due to moral, societal, or personal reasons, evoke feelings of discomfort, offense, or outrage when discussed or exposed in public settings.

To successfully execute the adversary model, the assumption is that the adversary objective is to communicate with the voice assistant to access and retrieve private information, which could be done in-person or remotely overt the Internet. For example, in-person may occur where the VA device is positioned anywhere within the user's house. The adversary, which may include family gatherings, parties, and other individuals, can freely walk in the house. The adversary can issue voice commands to the VA and intercept its responses.

Remotely over the internet such as in a Zoom meeting call, it is assumed the VA device is placed in close proximity to another device, such as a laptop or PC, that is being used the Zoom meeting [87], [88]. In this case, the device can intercepts the adversary's voice commands and provide responses.

Understanding the need to test for sensitive information within Amazon Alexa shopping service can hold sensitive details related to a user's shopping preferences, purchases, and even health or personal choices. This sensitivity arises from the nature of products or items that users may add to their shopping lists, carts, or purchase history, including medications, personal care products, or items that could be considered socially sensitive. The potential for privacy breaches or unauthorized access to such information, including purchase history, underscores the necessity of thorough testing.

In order to test the Amazon Alexa shopping service, it is necessary to address primarily the challenge to automate the testing process effectively. This challenge stems from uncertainties about the voice assistant's ability to manage repetitive commands, the variability in Amazon product names, and the understanding of feedback responses. To address this overarching challenge, this work proposes the Testing Voice Interface System (TVOS) framework, which consists of collecting commands, generating commands, and analyzing responses. TVOS streamlines the testing process and enhances the evaluation of Amazon Alexa's shopping service, addressing the critical need for automated testing in this context.

IV. FRAMEWORK FOR TESTING VOICE INTERFACE

TVOS is a framework designed to facilitate the efficient and automated testing of voice interface systems' prompts when queries about the shopping service. TVOS operates by creating simulated voice interactions with the target voice interface system, such as Alexa, replicating user requests, and capturing responses. The framework is shown in Figure 2.

A. FRAMEWORK DESIGN

The framework addressed the challenge by introducing diversity in voice commands when issuing commands to the voice assistant, which is achieved by using various voices provided by a Text-to-Speech (TTS) service. The framework also incorporates a pause before repeating the same question to mimic natural user interactions. It comprises the following components: *Collection of voice commands and products*, *Generate commands*, and *Response Aggregator*.

1) COLLECTING VOICE COMMANDS AND PRODUCT

Voice commands and product play a vital role in the TVOS framework. They encompass the diverse range of questions and commands users typically issue to smart speakers for daily tasks. These user inputs are fundamental for conducting realistic and comprehensive tests. Likewise, the selection of product is essential, particularly when addressing sensitive information within voice-controlled ecosystems. These voice commands and product categories are essential for command generation.

2) GENERATE COMMANDS

The process of command generation accommodates for a variations in product names. The generation involves the substitution of the [product] in the predefined command templates with sensitive product titles to generate a set of test cases. Initially, command templates such as *"Do I have [product] in my cart"* are selected and [product] is substitute with titles from the predefined Amazon product list. Product titles can vary significantly in length. In the first round, the complete product title as listed on Amazon is used. When the voice assistant initially indicates an item's absence in the shopping cart, variations in the product title are introduced. For instance, if a product title is extensive, such as *"VCF Vaginal Contraceptive Film With Spermicide, 5 Boxes of 9 Prevents Pregnancy, Nonoxynol-9 Kills Sperm on Contact, Hormone-Free, Easy to Use, Unnoticeable, 45 Total,"* and the VA responds with *"You don't have this item in the cart,"* the product title is shortened. The title is iteratively refined by removing portions after the first comma (",") or hyphen ("-") or eliminating prepositions such as "for," "with," "by," and others. To streamline and automate this process, a Python script is utilized, allowing efficient test case generation and systematic assessment of the VA's responses under varying product title conditions.

3) RESPONSE AGGREGATOR

After the command is sent to the voice assistant, it provides a feedback. The framework should be able to understand these responses in order to evaluate when the VA prompts products. To understand diverse responses, the responses were classified into different types, and further analyze usability according to their types. The VA responses can be classified into two types: failure response and recognized product response. If the VA fails to recognize the command, it will provide an error message such as *"sorry, I don't understand that"*. If the VA recognize a product in the input command such as *"Do I have [product] in my shopping cart?*, it will reply *"Yes, I found [product] in your Amazon cart"* indicating that VA recognize the user request and disclose the information. The rules are summarized as shown in Table 3.

B. VALIDATION PROCESS

The validation process involves a rule-based approach to understand voice assistant responses, develop response databases, and systematically analyze and label the responses. A rule-based approach is developed since the VA's responses have clear patterns, such as using the words *"sorry,.."* or *"My apologies.."* as for failure responses. The reason is that the samples is provided in Amazon documentation and guideline for end users, for example in case of failure or unaccepted commands the VA responses

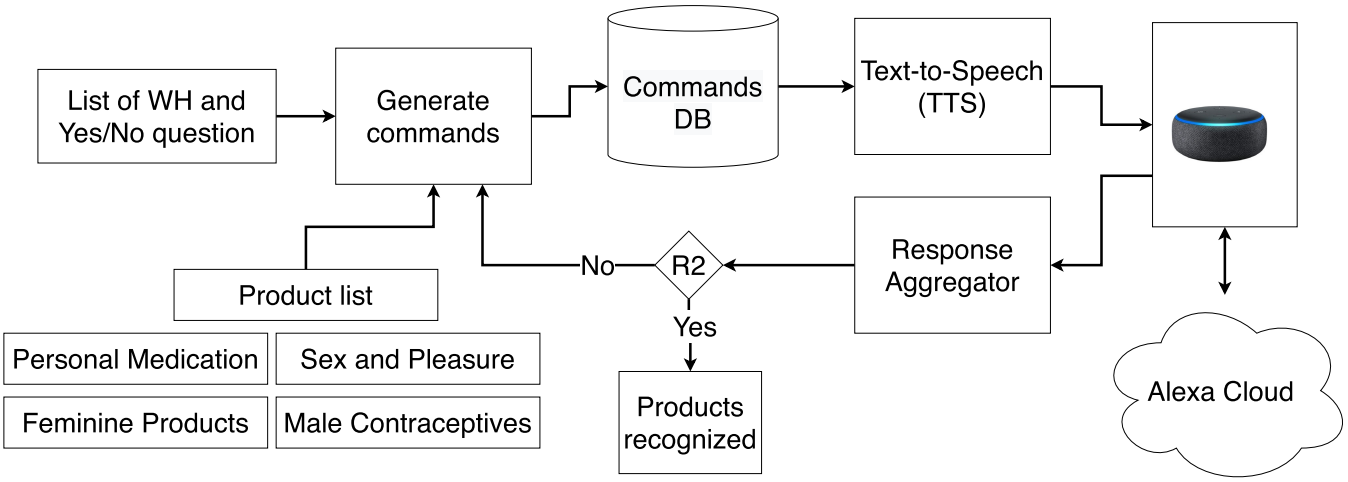

FIGURE 2. TVOS Framework Interaction: Illustrating the flow of voice interactions within the TVOS framework, capturing the queries and responses with the target voice interface system.

TABLE 3. Rules and response classifications for understanding user queries.

Rules	Situation	Response Classification
Rule 1	"sorry, I don't understand", "My apologies, please repeat your command", "I am having trouble understanding", "I don't get that"	Failure
Rule 2	statement, "yes, I found [product] by XXX in your Amazon cart", "you have [Number], [product1], [product2] in your cart/list", "You ordered [product] for $, it was ordered on YYYY,MM,DD"	Recognized Products

"sorry, I don't understand" [89]. In order to create a database on VA responses, interactions with the VA are conducted and the voice assistant responses are collected. The responses are then manually analyzed, labeled, and compiled into a response database related to the response classification. Subsequently, matching rules derived from the guidelines outlined in Table 3 are established to determine which pattern to employ in crafting the response. This implementation utilizes a Python library designed for rule-based matching in Natural Language Processing (NLP). The process begins by defining the specific patterns aimed to identified. These patterns are then added into the Matcher tool. Subsequently, the Matcher tool is applied to the responses generated by the interaction tool, referred to as the Response Aggregator. Finally, the Matcher systematically scans the knowledge base for the encoded patterns. It proceeds word by word, initially lemmatizing each word to its base form, contextualizing it according to its meaning (e.g., "be" serves as the lemma for both "was" and "is"). If Rule 2 is captured, the corresponding command is added to the list of products exposed; otherwise, another product title length is considered or another command is selected for testing.

The relevance of responses within the framework, in the context of access control, can distinguish between Rule 1 and Rule 2 responses. Rule 2 responses serve as indicators of the system granting access to specific shopping data, while Rule 1 responses signify instances where access is withheld due to complications arising from command misunderstandings or the deliberate prevention of unauthorized access as a response to ambiguous or inappropriate queries.

TABLE 4. Simple commands.

Category	Command Template
Cart	"What's in my shopping cart?", "What's on my cart?", "show me my cart?", "show my shopping cart?"
List	"What is on my shopping list?", "Show me my shopping list?", "what's on my list?"
History	"What did I buy last week/month/year?", "what is my order history?", "Show me my recent purchase history?"

V. INTERACTION STRATEGIES AND RESULTS

This section details the strategies an adversary might leverage to exploit vulnerabilities and circumvent the access control measures implemented by Amazon Alexa, thereby gaining unauthorized access to a user's private information.

Dataset Collection: To facilitate testing, a comprehensive list of voice commands that users might employ, along with a list of Amazon products collected from the Amazon marketplace, are populated.

Voice Commands: An investigation was conducted into potential questions and commands that users might pose to smart speakers for daily tasks [90], [91], [92], [93]. These questions and commands, particularly those related to shopping such as *"What is my shopping cart/list?"* or *"Add milk in shopping cart/list?"* were collected. This collection was conducted manually due to the small size of the sets. Table 4 shows a list of possible commands that users can use when interacting with smart home assistants in general when shopping, preparing for shopping, or querying the shopping history.

TABLE 5. Product categories and definitions.

Category	Definition and Examples
Male Contraceptives	Male contraceptives are used to prevent pregnancy and sexually transmitted infections (e.g., condoms).
Feminine Products	Personal care products used during menstruation, such as pads. Also, this includes female contraceptives (e.g., birth control pills and hormonal contraceptives).
Personal Medication	Medications for various health conditions, including depression, anxiety, erectile dysfunction, male enhancement, sexual diseases like STDs (Sexually Transmitted Diseases).
Sex and Pleasure	Products and items related to sexual well-being and pleasure (e.g., sex toys, lubricants, etc.).

TABLE 6. Dataset statistics: Amazon products and purchases history over four years.

Year	Amazon Products	Products Purchase History
2020	-	17
2021	-	23
2022	-	28
2023	80	20

Amazon Products: Regarding products that users would use the smart home assistant to shop for, investigation was conducted into what are termed as sensitive products or controversial products. This does not refer to illegal products, but rather to items that individuals might find unsavory, such as personal medication, feminine hygiene products, tobacco, and birth control. Line of research in this area has described these products as:"unmentionables," "socially sensitive products," or "controversial products" [38], [39], [40], [41], [85], [86]. They defined these products or concepts that for the reasons of morality or fear tend to elicit a reaction of disgust, offense or outrage when mentioned or presented in public. This study categorizes these private sensitive products into four categories based on previous studies; refer to Table 5.

This work employed a selenium-based web crawler [94], to access the Amazon marketplace and conduct searches for product categories listed in Table 5. The web crawler facilitated the retrieval of comprehensive product information, including titles, ratings, brands, and other relevant details. To ensure data quality and relevance, this work focused on gathering products within specific categories, sorting them by user ratings, and selecting the top 20 products from each category. Consequently, the dataset comprises a total of 80 Amazon products, all of which were collected from the Amazon U.S. store in September 2023. Additionally, an extensive purchase history maintained by the author since 2020 was leveraged. The purchase history was scraped to get list of these products along with product detail information. The purchase history provided a dataset of products, including detailed attributes such as product title, brand, and form, which were selected based on their frequency of occurrence within the author's order history; see Table 6 for dataset statistics.

A. FLAW 1: USING SIMPLE COMMAND

1) TESTING PROCEDURE

In this experiment, the adversary interacts with the Echo device to access and retrieve private information by employing simple commands that mimic common user queries. Table 4 shows a list of simple commands the adversary can use to perform, which query for the content of the shopping cart, list, and history.

The experiment was conducted to account for retrieval information. For each product category outlined in Table 5, a systematic approach was implemented to add the collected products to both the Amazon shopping cart and shopping list. A script was created to automate interactions with the Amazon marketplace, adding products to the cart. To integrate the selected products into the Alexa shopping list, Amazon Alexa's user-friendly online shopping list interface was utilized [95]. This process involved tasks like adding, deleting, and editing items in the shopping list. To assess shopping history across various time-frames, access was granted to the author's purchase history dating back to 2020. Insights into product orders spanning weeks, months, and years were gained by relying on past orders from the author's Amazon account.

2) TESTING SETUP

For speech synthesis and recognition, the pyttsx3 library was utilized for text-to-speech conversion, which offers a variety of voice IDs [84]. The PyPI library was employed for speech recognition [96]. A rule-based approach was used, employing SpaCy, an open-source software library for advanced NLP, to implement matching rules. All experiments were conducted using TVOS framework (see Figure 2) on a third-generation Echo device and an Apple MacBook with an Apple M1 Pro processor and 32GB of RAM.

Text to Speech (TTS) is responsible for automating the delivery of a command to the smart speaker. It takes a text command and converts it into a synthesized voice, which is then transmitted to the Alexa Echo device. The Echo device provides an audio response, which is captured by the response aggregator. The response aggregator records the Echo's responses and converts them into text using the speech recognition module. The response aggregator communicates with the Google API to convert speech to text and generates a file containing responses.

3) RESULTS

Table 7 presents the results of using the simple commands to access user's private shopping information. The findings indicate that the Alexa voice assistant grants access to shopping data by listing *all* the products in the shopping

TABLE 7. Results of using the simple commands to expose user's private information. *All* indicates that the VA exposes all the products/items in the shopping cart or list. *Partial* means the VA will list a few products purchased in the last 30 days.

(a) Shopping Cart and List

Product Category	Cart	List
Male Contraceptives	All	All
Feminine Products	All	All
Personal Medication	All	All
Sex and Pleasure	All	All

(b) Purchase History

Year	History
2020	0
2021	0
2022	0
2023	Partial

cart and shopping list. When queried about the shopping cart, Alexa provides a list of items in the cart and asks if the user would like to hear the rest. For example, if the cart contains 20 items, Alexa responds with, *"You have 20 items in your Amazon cart, including [product1], [product2], and [product3]. Would you like to hear the rest?"* Similarly, if these 20 items are added to Alexa's shopping list, it responds with, *"You have 20 items on your shopping list. Here are the 5 most recent: [item1], [item2], [item3], [item4], and [item5]. Would you like to hear the next 5 items?"* (see Table 7a).

When queried about purchase history, Alexa provides *partial* lists of items ordered in the last 30 days, such as *"In the last 30 days, you've ordered X items, including [item1], [item2], and so on."* Regardless of the specific commands used from the simple commands, Alexa consistently provides access to items purchased in the last 30 days, exposing the most recent purchases, such as the 5 items ordered by the author within this timeframe (see Table 7b). Therefore, the current access control measures implemented by Alexa for the shopping cart and shopping list are inadequate, as they grant access to all contents, exposing the titles of each product. In contrast, regarding shopping history protection, Alexa does not provide full access to historical data, which is a relatively better safeguard than the shopping cart and shopping list.

Takeaways: The exploration of flaw 1 reveals the inadequacy of protection for user-sensitive information, which allows the obtaining of private information about a user's shopping cart and shopping list. However, shopping history is protected with no access for more than a month time-frame.

B. FLAW 2: USING TARGETED COMMAND

Based on results from the initial exploration of system vulnerabilities (Flaw 1), which involved querying for product information, a targeted approach was implemented. This approach leverages knowledge of specific product details, such as titles, brands, forms, and active ingredients, which can be readily obtained from sources like Amazon market. This targeted exploration is motivated by two primary objectives: A) While the simple commands (Flaw 1) provided a limited access to the purchase history, the primary objective is to investigate the potential for accessing the entire user purchase history. B) Additionally, while issuing simple commands explores all items within the shopping cart and list, the adversary might obtain unrelated items that are of no interest. The adversary is primarily concerned with sensitive and privacy-related items. The targeted exploration aim to explore the potential of targeted command to expedited access to information.

TABLE 8. Targeted commands.

Category	Command Template
Cart	"Check my shopping cart, do I have [product title/brand/form]?" , "Do I have [product title/brand/form] in my shopping cart?" , "Do I have [product title/brand/form] in my cart?", "Is [product title/brand/form] in my Cart?"
List	"Is [product title] in my shopping list?", "Do I have [product title] in my shopping list?"
History	"When is the last time I bought [product title/brand/form]?" , "Did I buy [product title/brand/form] in the past?", "Did I order [product title/brand/form] in the past?"

Table 8 provides a list of targeted commands that the adversary can employ to acquire private information about specific products. These commands, such as *"When is the last time I bought [product title/brand/form]?"* or *"Check my shopping cart, do I have [product title/brand/form]?"* are designed to extract precise data about individual products, utilizing various product attributes. By implementing these targeted commands, the adversary aims to gain access to detailed information about products that the user may consider private information.

1) TESTING PROCEDURE

To further evaluate the effectiveness of the targeted command to access the user purchase history, a purchase history dataset maintained by the author since 2020 was leveraged. In regard of the shopping cart and list, the same data was utilized as before by adding the products to the shopping cart and list. It is worth noting that the cart and list serve distinct purposes within Amazon's ecosystem. Users utilize the shopping cart to add products from Amazon.com, and this feature includes detailed product information, such as title, brand and others. The shopping list, on the other hand, is designed for users to compile lists of items they intend to purchase. In this experiment, all items added to the shopping list were treated as equivalent to product titles. The same testing setup as in the exploration of Flaw 1 is utilized.

2) RESULTS

The targeted commands were employed to retrieve private information from the user's purchase history, and the results are presented in Figure 3. The chart displays success rates for each feature, where higher percentages indicate more successful retrievals. Notably, utilizing the product *title* and *brand* allowed the adversary to retrieve the entire purchase history from 2020 to 2023 with a success rate of 100%. However, for the product feature like *form*, the capabilities were limited to capturing partial product data from the purchase history, achieving a success rate between 26%

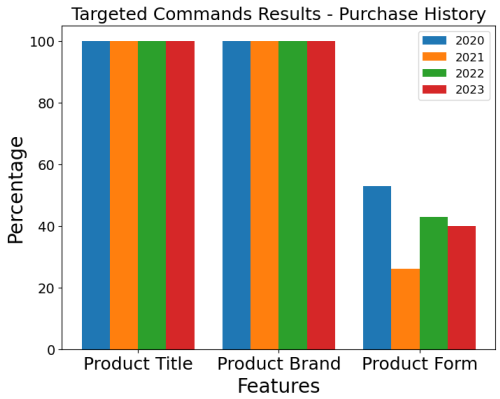

FIGURE 3. Percentage success rates of targeted commands to expose user private information on the purchase history. The data is categorized based on different product features over four years shopping history.

and 52%. Notably, Alexa provided purchase information for product *form* such as "oil," "lotion," "drops," "spray," and "cream," but remained non-responsive to queries concerning other form, including "powder," "tablet," and "capsule."

In cases where a product was ordered multiple times (e.g., in both 2020 and 2022), Amazon Alexa consistently favored the most recent purchase date (2022) in its responses. Even when the targeted command was modified with specific product titles or brands and the exact date of a previous purchase (e.g., 2020), Alexa's responses continued to prioritize the recent purchase date (2022). In instances where the product was ordered only once, the targeted command consistently revealed purchase information dating back to 2020, eliciting responses such as *"You ordered [product title] for $$. It was ordered on YYYY/MM/DD., Currently, it's $$, If you need more, just say buy it now."*

Figure 4 provides a visual representation of the outcomes from the targeted commands, aimed to expedite the retrieval and access of private information from both the shopping cart and shopping list, offering a detailed perspective on the success rates for retrieving private information. The data has been categorized into different product features and showcases the success rates for each category. The heatmap demonstrates success rates in the shopping cart, with higher percentages indicating more successful retrievals. Notably, the product *brand* consistently provided comprehensive access to desired shopping data. When utilizing the specified product feature, comprehensive access to relevant product information was successfully retrieved from the shopping cart. While not all products were accessible using the product *title* and *form*, this approach still provided valuable insights into a subset of the products from the shopping cart. This was the case with feminine and personal medication products, with a success rate between 35% and 55%. For products from male contraceptive and sex and pleasure categories, the features *title* and *form* were ineffective for accessing the desired data. This ineffectiveness is believed to be due to the *title* and *form* containing content that could be considered inappropriate or sensitive, such as condoms and adult toys. While Amazon Alexa's internal information

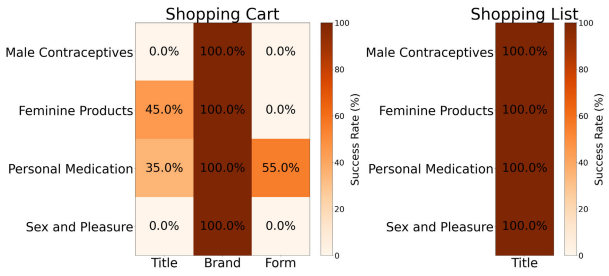

FIGURE 4. Percentage success rates of targeted commands for expedited retrieval from the shopping cart and shopping list.

regarding the implementation of a profanity filter are not publicly disclosed, the response from Alexa, *"Sorry, I don't have an answer for that."* suggests that the recognition system may be filtering them out.

In the shopping list results, presented in the right heatmap, success rates for each category are visible, demonstrating the effectiveness of targeted commands to access private product information from the shopping list. Utilizing the product *title* allowed us to list all items on the shopping list. When using the product *title* feature, Alexa lists all the items on the shopping list and responds with *"You have 20 items on your shopping list. Here are the 5 most recent: [item1], [item2], [item3], [item4], and [item5]. Would you like to hear the next 5 items?"*. However, it would not provide specific product details related to the queried title. Instead, it presented a comprehensive list of items without offering individualized information.

Takeaways: The exploration of targeted commands (Flaw 2) yielded varying outcomes across different product categories and features. Product ***title*** and ***brand*** proved highly effective, granting access to the user's complete purchase history. However, certain features like ***form*** only allowed partial access to shopping history. Results varied when retrieving data from the shopping cart and list. Intriguingly, the ***brand*** feature provided a workaround, bypassing filtering mechanisms to access the user's shopping information from the shopping cart. For features such as ***title***, partial retrieval of products was achieved. The sensitivity of product titles, especially in the "male contraceptives" and "sex and pleasure" categories, led to blocked access.

VI. RECOMMENDATIONS

This section presents recommendations aimed at enhancing the security measures that safeguard user private data and addressing concerns related to unauthorized access. The recommendations are organized as follows:

A. BETTER CONTROL OVER USER DATA

Implementing personalized notifications for sensitive actions is recommended to enhance user privacy and security over user data. Users should be allowed to identify what parts of the voice service are accessible and only sent to the user's phone. For example, in the shopping app, the user can identify either a shopping cart, shopping list, or shopping

history component to be sent to the primary account holder's smartphone or other designated device, prompting user confirmation before allowing the action to proceed [97], [98], [99], [100]. Thus, the user data will only be available through visual interfaces, and the voice assistants will not audibly disclose the action (data). This approach can protect and safeguard user's private data by requiring user validation, ensuring that potentially sensitive information is disclosed only with the user's explicit consent.

B. BETTER CONTROL OVER DEVICE FUNCTIONALITY AND CONFIGURABILITY

The voice assistant services should have privacy modes that users can enable and manage access to the data such as shopping data [14], [46], [75], [101], [102]. With privacy modes such as family mode or guest mode, the user can specify what he/she means by family mode, such as always making everything private or partial data private but not the other when the family or guest members are present. The mode configurability can be accomplished on the user's phone or web portal, where the user can specify the definition of the mode. For example, if the user configures the family mode to always make everything private and activates the mode, it temporarily disables access to the user data, such as shopping data. Thus, it ensures that the voice assistant service functionality on the shopping data, for example, remains inactive until reactivated by the user. The privacy mode control is more usable than a mute button on the smart speaker, and it is practical as users can enable, disable, and customize it to meet their individual needs to disable access to private data such as historical shopping data but not the shopping list.

VII. CONCLUSION

Access control is fundamental to ensuring the privacy and security of voice-activated systems. Thus study delved into examining access control mechanisms for voice-controlled systems in multi-user environments. Critical vulnerabilities were revealed in these access control systems, highlighting the pressing need for more robust privacy protections. The analysis highlights two flaws within these access control systems. The first flaw, where an adversary exploits simple commands to access and retrieve private information from the user's shopping cart, list, and purchase history, a clear inadequacy was observed in protecting user-sensitive information. This vulnerability allowed unauthorized access to a user's shopping cart and shopping list, with shopping history being only partially protected, exposing the most recent purchases within a month. The second flaw, which involves utilizing targeted commands for expedited retrieval of private information from the shopping cart and shopping list, yielded varying outcomes across product categories and features. Notably, using product titles and brands proved highly effective, granting access to the user's complete purchase history. The implications of this study are significant for user privacy and security, underscoring the importance of implementing robust measures to prevent unauthorized access to sensitive information. This work have provided recommendations to enhance security measures and address concerns related to unauthorized access. While the investigation primarily focused on Amazon Alexa, the broader landscape of voice user interface platforms where users interact with voice assistants for shopping and similar functionalities is acknowledged. It is encouraged for service providers to invest in more robust access control systems. Future research should continue exploring and addressing access control challenges in diverse voice-controlled systems, fostering continuous improvement in privacy and security measures.

REFERENCES

[1] T. Bergur. *Smart Home Statistics*. Accessed: Oct. 2023. [Online]. Available: https://www.statista.com/topics/2430/smart-homes/#topicOverview
[2] J. Koetsier. (2018). *Amazon Echo, Google Home Installed Base Hits 50 Million; Apple Has 6% Market Share, Report Says*. [Online]. Available: https://www.forbes.com/sites/johnkoetsier/2018/08/02/amazon-echo-google-home-installed-base-hits-50-million-apple-has-6-market-share-report-says
[3] Google. *Google Home Product Page*. Accessed: Sep. 2023. [Online]. Available: https://store.google.com/product/google_home
[4] Amazon. *Alexa Store*. Accessed: Jul. 2023. [Online]. Available: https://www.amazon.com/alexa-skills
[5] E. Schwartz. (2023). *Alexa-Enabled Device Sales Pass 500m As Amazon Unveils New and Upgraded Echo Devices*. Accessed: May 17, 2023. [Online]. Available: https://voicebot.ai/2023/05/17/
[6] N. Abdi, K. M. Ramokapane, and J. M. Such, "More than smart speakers: Security and privacy perceptions of smart home personal assistants," in *Proc. 15th Symp. Usable Privacy Secur.*, Santa Clara, CA, USA, Aug. 2019, pp. 451–466. [Online]. Available: https://www.usenix.org/conference/soups2019/presentation/abdi
[7] A. Sciuto, A. Saini, J. Forlizzi, and J. I. Hong, "'Hey Alexa, what's up?' A mixed-methods studies of in-home conversational agent usage," in *Proc. Designing Interact. Syst. Conf.*, 2018, pp. 857–868.
[8] J. S. Edu, J. M. Such, and G. Suarez-Tangil, "Smart home personal assistants: A security and privacy review," *ACM Comput. Surv.*, vol. 53, no. 6, pp. 1–36, Nov. 2021.
[9] T. Ammari, J. Kaye, J. Y. Tsai, and F. Bentley, "Music, search, and IoT: How people (really) use voice assistants," *ACM Trans. Computer-Human Interact.*, vol. 26, no. 3, pp. 1–28, Jun. 2019.
[10] A. Tilley. (2016). *How a Few Words To Apple's Siri Unlocked a Man's Front Door*. [Online]. Available: http://www.forbes.com/sites/aarontilley/2016/09/21/apple-homekit-siri-security
[11] T. Huxohl, M. Pohling, B. Carlmeyer, B. Wrede, and T. Hermann, "Interaction guidelines for personal voice assistants in smart homes," in *Proc. Int. Conf. Speech Technol. Human-Computer Dialogue (SpeD)*, Oct. 2019, pp. 1–10.
[12] F. K. Santoso and N. C. H. Vun, "Securing IoT for smart home system," in *Proc. Int. Symp. Consum. Electron. (ISCE)*, Jun. 2015, pp. 1–2.
[13] E. Zeng and F. Roesner, "Understanding and improving security and privacy in multi-user smart homes: A design exploration and in-home user study," in *Proc. 28th USENIX Secur. Symp.*, 2019, pp. 159–176.
[14] J. Lau, B. Zimmerman, and F. Schaub, "Alexa, are you listening: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers," *Proc. ACM Hum.-Comput. Interact.*, vol. 2, pp. 1–31, Nov. 2018.
[15] C. Tsiourti, M. Ben Moussa, J. Quintas, B. Loke, I. Jochem, J. A. Lopes, and D. Konstantas, "A virtual assistive companion for older adults: Design implications for a real-world application," in *Proc. SAI Intell. Syst. Conf.*, vol. 1. Cham, Switzerland: Springer, 2018, pp. 1014–1033.
[16] A. K. Sikder, L. Babun, Z. B. Celik, H. Aksu, P. McDaniel, E. Kirda, and A. S. Uluagac, "Who's controlling my device? Multi-user multi-device-aware access control system for shared smart home environment," *ACM Trans. Internet Things*, vol. 3, no. 4, pp. 1–39, Nov. 2022.

[17] M. Cecchinato and D. Harrison, "Degrees of agency in owners and users of home IoT devices," in *Proc. Workshop, Making Home, Asserting Agency Age IoT*, 2017, pp. 1–5.

[18] E. Zeng, S. Mare, and F. Roesner, "End user security and privacy concerns with smart homes," in *Proc. 13th Symp. Usable Privacy Security (SOUPS)*, Jul. 2017, pp. 65–80.

[19] Y. Huang, B. Obada-Obieh, and K. Beznosov, "Amazon vs. my brother: How users of shared smart speakers perceive and cope with privacy risks," in *Proc. CHI Conf. Human Factors Comput. Syst.*, Apr. 2020, pp. 1–13.

[20] K. Marky, S. Prange, M. Mühlhäuser, and F. Alt, "Roles matter! Understanding differences in the privacy mental models of smart home visitors and residents," in *Proc. 20th Int. Conf. Mobile Ubiquitous Multimedia*, May 2021, pp. 108–122.

[21] K. Marky, A. Voit, A. Stöver, K. Kunze, S. Schröder, and M. Mühlhäuser, "'I don't know how to protect myself': Understanding privacy perceptions resulting from the presence of bystanders in smart environments," in *Proc. 11th Nordic Conf. Human-Comput. Interact., Shaping Experiences, Shaping Soc.*, 2020, pp. 1–11.

[22] H. Yoon, S.-H. Park, and K.-T. Lee, "Lightful user interaction on smart wearables," *Pers. Ubiquitous Comput.*, vol. 20, no. 6, pp. 973–984, Nov. 2016.

[23] A. Renz, M. Baldauf, E. Maier, and F. Alt, "Alexa, it's me! An online survey on the user experience of smart speaker authentication," in *Proc. Mensch Und Comput.*, Sep. 2022, pp. 14–24.

[24] A. Renz, T. Neff, M. Baldauf, and E. Maier, "Authentication methods for voice services on smart speakers—A multi-method study on perceived security and ease of use," *I-COM*, vol. 22, no. 1, pp. 67–81, Apr. 2023.

[25] Z. Meng, M. U. B. Altaf, and B.-H. Juang, "Active voice authentication," *Digit. Signal Process.*, vol. 101, Jun. 2020, Art. no. 102672.

[26] X. Wang, Z. Yan, R. Zhang, and P. Zhang, "Attacks and defenses in user authentication systems: A survey," *J. Netw. Comput. Appl.*, vol. 188, Aug. 2021, Art. no. 103080.

[27] Z. Yan and S. Zhao, "A usable authentication system based on personal voice challenge," in *Proc. Int. Conf. Adv. Cloud Big Data (CBD)*, Aug. 2016, pp. 194–199.

[28] J. B. Horrigan, "Online shopping," Pew Internet Amer. Life Project, Washington, DC, USA, 2008.

[29] F. Kawsar and A. J. B. Brush, "Home computing unplugged: Why, where and when people use different connected devices at home," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Sep. 2013, pp. 627–636.

[30] A. M. Alrabei, L. N. Al-Othman, F. A. Al-Dalabih, T. A. Taber, and B. J. Ali, "The impact of mobile payment on the financial inclusion rates," *Inf. Sci. Lett.*, vol. 11, no. 4, pp. 1033–1044, 2022.

[31] D. N. Aryani, R. K. Nair, D. X. Y. Hoo, D. K. M. Hung, D. H. R. Lim, D. A. R. Chandran, W. P. Chew, and A. Desai, "A study on consumer behaviour: Transition from traditional shopping to online shopping during the COVID-19 pandemic," *Int. J. Appl. Bus. Int. Manag.*, vol. 6, no. 2, pp. 81–95, Aug. 2021.

[32] D. Grewal, M. Levy, and V. Kumar, "Customer experience management in retailing: An organizing framework," *J. Retailing*, vol. 85, no. 1, pp. 1–14, Mar. 2009.

[33] V. Rabassa, O. Sabri, and C. Spaletta, "Conversational commerce: Do biased choices offered by voice assistants' technology constrain its appropriation?" *Technol. Forecasting Social Change*, vol. 174, Jan. 2022, Art. no. 121292.

[34] K. Bret and M. Ava. (2018). *Voice Shopping Consumer Adoption Report*. Accessed: Aug. 2023. [Online]. Available: https://voicebot.ai/wp-content/uploads/2018/06/voice-shopping-consumer-adoption

[35] R. E. Bawack, S. F. Wamba, and K. D. A. Carillo, "Adoption of smart speakers for voice shopping," in *Exploring Innovation in a Digital World: Cultural and Organizational Challenges*. Berlin, Germany: Springer, 2021, pp. 21–35.

[36] A. A. Salameh, I. A. Abu AlSondos, B. J. A. Ali, and A. F. Alsahali, "From citizens overview: Which antecedents' can assist to increase their satisfaction towards the ubiquity of mobile commerce applications?" *Int. J. Interact. Mobile Technol. (iJIM)*, vol. 14, no. 17, p. 45, Oct. 2020.

[37] K.-N. Wu, "Voice assistant," School Comput. Sci., Univ. Birmingham, Birmingham, U.K., Tech. Rep., 2018.

[38] A. Wilson and C. West, "The marketing of unmentionables," *Harvard Bus. Rev.*, vol. 59, no. 1, pp. 91–102, 1981.

[39] D. W. Dahl, R. V. Manchanda, and J. J. Argo, "Embarrassment in consumer purchase: The roles of social presence and purchase familiarity: Table 1," *J. Consum. Res.*, vol. 28, no. 3, pp. 473–481, Dec. 2001.

[40] A. Krishna, K. B. Herd, and N. Z. Aydinoglu, "A review of consumer embarrassment as a public and private emotion," *J. Consum. Psychol.*, vol. 29, no. 3, pp. 492–516, Jul. 2019.

[41] D. S. Waller, K. Fam, and B. Zafer Erdogan, "Advertising of controversial products: A cross-cultural study," *J. Consum. Marketing*, vol. 22, no. 1, pp. 6–13, Jan. 2005.

[42] S. Virupaksha, D. Gavini, and D. Venkatesulu, "Data privacy in online shopping," in*Computer Communication, Networking and Internet Security*. Singapore: Springer, 2017, pp. 199–207.

[43] B. Kinsella, "Amazon Alexa has 100K skills but momentum slows globally. Here is the breakdown by country. Voicebot. AI," Voicebot.ai, Washington, DC, USA, Tech. Rep., 2020.

[44] D. J. Dubois, R. Kolcun, A. M. Mandalari, M. T. Paracha, D. Choffnes, and H. Haddadi, "When speakers are all ears: Characterizing misactivations of IoT smart speakers," *Proc. Privacy Enhancing Technol.*, vol. 2020, no. 4, pp. 255–276, Oct. 2020.

[45] W. He, M. Golla, R. Padhi, J. Ofek, M. Dürmuth, E. Fernandes, and B. Ur, "Rethinking access control and authentication for the home Internet of Things," in *Proc. 27th USENIX Secur. Symp.*, 2018, pp. 255–272.

[46] L. Schönherr, M. Golla, T. Eisenhofer, J. Wiele, D. Kolossa, and T. Holz, "Unacceptable, where is my privacy? Exploring accidental triggers of smart speakers," 2020, *arXiv:2008.00508*.

[47] X. Yuan, Y. Chen, and Y. Zhao, "CommanderSong: A systematic approach for practical adversarial voice recognition," in *Proc. USENIX Secur. Symp.*, 2018, pp. 49–64.

[48] N. Abdi, X. Zhan, K. M. Ramokapane, and J. Such, "Privacy norms for smart home personal assistants," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2021, pp. 1–14.

[49] A. Cosson, A. K. Sikder, L. Babun, Z. B. Celik, P. McDaniel, and A. S. Uluagac, "Sentinel: A robust intrusion detection system for IoT networks using kernel-level system information," in *Proc. Int. Conf. Internet-Things Design Implement.*, May 2021, pp. 53–66.

[50] S. Demetriou, N. Zhang, Y. Lee, X. Wang, C. A. Gunter, X. Zhou, and M. Grace, "HanGuard: SDN-driven protection of smart home WiFi devices from malicious mobile apps," in *Proc. 10th ACM Conf. Secur. Privacy Wireless Mobile Netw.*, Jul. 2017, pp. 122–133.

[51] E. Fernandes, J. Paupore, A. Rahmati, D. Simionato, M. Conti, and A. Prakash, "FlowFence: Practical data protection for emerging IoT application frameworks," in *Proc. USENIX Secur.*, 2016, pp. 531–548.

[52] I. Agadakos, P. Hallgren, D. Damopoulos, A. Sabelfeld, and G. Portokalidis, "Location-enhanced authentication using the IoT: Because you cannot be in two places at once," in *Proc. 32nd Annu. Conf. Comput. Secur. Appl.*, Dec. 2016, pp. 251–264.

[53] S. Cirani, M. Picone, P. Gonizzi, L. Veltri, and G. Ferrari, "IoT-OAS: An OAuth-based authorization service architecture for secure services in IoT scenarios," *IEEE Sensors J.*, vol. 15, no. 2, pp. 1224–1234, Feb. 2015.

[54] M. Jacobs, H. Cramer, and L. Barkhuus, "Caring about sharing: Couples' practices in single user device access," in *Proc. 19th Int. Conf. Supporting Group Work*, Nov. 2016, pp. 235–243.

[55] A. I. Newaz, A. K. Sikder, L. Babun, and A. S. Uluagac, "HEKA: A novel intrusion detection system for attacks to personal medical devices," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Jun. 2020, pp. 1–9.

[56] H. Ren, Y. Song, S. Yang, and F. Situ, "Secure smart home: A voiceprint and internet based authentication system for remote accessing," in *Proc. 11th Int. Conf. Comput. Sci. Educ. (ICCSE)*, Aug. 2016, pp. 247–251.

[57] T. Matthews, K. Liao, A. Turner, M. Berkovich, R. Reeder, and S. Consolvo, "'She'll just grab any device that's closer' a study of everyday device & account sharing in households," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2016, pp. 5921–5932.

[58] R. Garg and C. Moreno, "Understanding motivators, constraints, and practices of sharing Internet of Things," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 2, pp. 1–21, Jun. 2019.

[59] N. Carlini and D. Wagner, "Audio adversarial examples: Targeted attacks on speech-to-text," in *Proc. IEEE Secur. Privacy Workshops (SPW)*, May 2018, pp. 1–7.

[60] N. Carlini, P. Mishra, T. Vaidya, Y. Zhang, M. Sherr, C. Shields, D. Wagner, and W. Zhou, "Hidden voice commands," in *Proc. 25th USENIX Security Symp.*, Aug. 2016, pp. 513–530.

[61] G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu, "DolphinAttack: Inaudible voice commands," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2017, pp. 103–117.

[62] L. Schönherr, M. Golla, T. Eisenhofer, J. Wiele, D. Kolossa, and T. Holz, "Exploring accidental triggers of smart speakers," *Comput. Speech Lang.*, vol. 73, May 2022, Art. no. 101328.

[63] T. Le, D. Y. Huang, N. Apthorpe, and Y. Tian, "Skillbot: Identifying risky content for children in Alexa skills," *ACM Trans. Internet Technol. (TOIT)*, vol. 22, no. 3, pp. 1–31, 2021.
[64] Z. Guo, Z. Lin, P. Li, and K. Chen, "SkillExplorer: Understanding the behavior of skills in large scale," in *Proc. 29th USENIX Secur. Symp.*, 2020, pp. 2649–2666.
[65] J. Young, S. Liao, L. Cheng, H. Hu, and H. Deng, "SkillDetective: Automated policy-violation detection of voice assistant applications in the wild," in *Proc. USENIX Secur. Symp.*, 2022, pp. 1–19.
[66] H. A. Shafei and C. C. Tan, "Do smart speaker skills support diverse audiences?" *Pervas. Mobile Comput.*, vol. 87, Dec. 2022, Art. no. 101716.
[67] A. Mhaidli, M. K. Venkatesh, Y. Zou, and F. Schaub, "Listen only when spoken to: Interpersonal communication cues as smart speaker privacy controls," *Proc. Privacy Enhancing Technol.*, vol. 2020, no. 2, pp. 251–270, Apr. 2020.
[68] Y. Liu, A. Rahmati, Y. Huang, H. Jang, L. Zhong, Y. Zhang, and S. Zhang, "XShare: Supporting impromptu sharing of mobile phones," in *Proc. 7th Int. Conf. Mobile Syst., Appl., Services*, Jun. 2009, pp. 15–28.
[69] X. Ni, Z. Yang, X. Bai, A. C. Champion, and D. Xuan, "DiffUser: Differentiated user access control on smartphones," in *Proc. IEEE 6th Int. Conf. Mobile Adhoc Sensor Syst.*, Oct. 2009, pp. 1012–1017.
[70] R. Schuster, V. Shmatikov, and E. Tromer, "Situational access control in the Internet of Things," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2018, pp. 1056–1073.
[71] M. Yahyazadeh, P. Podder, E. Hoque, and O. Chowdhury, "Expat: Expectation-based policy analysis and enforcement for appified smart-home platforms," in *Proc. 24th ACM Symp. Access Control Models Technol.*, May 2019, pp. 61–72.
[72] Amazon. *Add and Manage Addresses*. Accessed: Sep. 2023. [Online]. Available: https://www.amazon.com/gp/help/customer/display.html?nodeId=GQT5HV6YYGNDSFNW
[73] Amazon. *Update Your Payment Method*. Accessed: Sep. 2023. [Online]. Available: https://www.amazon.com/gp/help/customer/display.html?nodeId=GPVBVQ28CKHD7ZQW
[74] Amazon. *Voice Code for Purchases*. Accessed: Sep. 2023. [Online]. Available: https://www.amazon.com/gp/help/customer/display.html?nodeId=GAA2RYUEDNT5ZSNK
[75] Amazon. *Alexa Voice Id*. Accessed: Sep. 2023. [Online]. Available: https://www.amazon.com/gp/help/customer/display.html?nodeId=GYCXKY2AB2QWZT2X
[76] Amazon. *Hange Your Device's, 'Wake Word'*. Accessed: Sep. 2023. [Online]. Available: https://www.aboutamazon.com/news/devices/tips-to-unlock-useful-and-fun-alexa-features
[77] E. D. Liddy, "Natural language processing," in *Encyclopedia of Library and Information Science*, 2nd ed. New York, NY, USA: Marcel Dekker, 2001.
[78] X. Lei, G.-H. Tu, A. X. Liu, K. Ali, C.-Y. Li, and T. Xie, "The insecurity of home digital voice assistants - Amazon Alexa as a case study," 2017, *arXiv:1712.03327*.
[79] B. D. Davis, J. C. Mason, and M. Anwar, "Vulnerability studies and security postures of IoT devices: A smart home case study," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10102–10110, Oct. 2020.
[80] A. Sabir, E. Lafontaine, and A. Das, "Hey Alexa, who am I talking to: Analyzing users' perception and awareness regarding third-party Alexa skills," in *Proc. CHI Conf. Human Factors Comput. Syst.*, Apr. 2022, pp. 1–15.
[81] Y. Chen et al., "Devil's whisper: A general approach for physical adversarial attacks against commercial black-box speech recognition devices," in *Proc. USENIX Secur.*, 2020, pp. 2667–2684.
[82] D. Kumar, R. Paccagnella, P. Murley, E. Hennenfent, J. Mason, A. Bates, and M. Bailey, "Skill squatting attacks on Amazon Alexa," in *Proc. 27th USENIX Secur. Symp.*, 2018, pp. 33–47.
[83] Apple. *Have Your Mac Speak Text*. Accessed: Oct. 2023. [Online]. Available: https://support.apple.com/en-gw/guide/mac-help/mh27448/mac
[84] *Text To Speech (TTS) Library*. Accessed: Oct. 2023. [Online]. Available: https://pypi.org/project/pyttsx3/
[85] D. S. Waller, "Consumer offense towards the advertising of some gender-related products," *J. Consum. Satisfaction, Dissatisfaction Complaining Behav.*, vol. 20, pp. 72–85, Jan. 2007.
[86] A. T. Shao and J. S. Hill, "Advertising sensitive products in magazines: Legal and social restrictions," *Multinational Bus. Rev.*, vol. 2, no. 2, pp. 16–24, 1994.
[87] S. Esposito, D. Sgandurra, and G. Bella, "ALEXA VERSUS ALEXA: Controlling smart speakers by self-issuing voice commands," in *Proc. ACM Asia Conf. Comput. Commun. Secur.*, May 2022, pp. 1064–1078.
[88] X. Yuan, Y. Chen, A. Wang, K. Chen, S. Zhang, H. Huang, and I. M. Molloy, "All your Alexa are belong to us: A remote voice control attack against echo," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.
[89] A. Pyae and P. Scifleet, "Investigating differences between native English and non-native English speakers in interacting with a voice user interface: A case of Google home," in *Proc. 30th Austral. Conf. Comput.-Human Interact.*, Dec. 2018, pp. 548–553.
[90] K. Bert. (2018). *What People Ask Their Smart Speakers*. Accessed: Aug. 2023. [Online]. Available: https://voicebot.ai/2018/08/01/what-people-ask-their-smart-speakers/
[91] *5 Ways Consumers Interact With Smart Speakers*. Accessed: Aug. 2023. [Online]. Available: https://mindstreammediagroup.com/
[92] Mindstreammediagroup. *Choosing the Right Smart Speaker for You: 5 Questions To Ask*. Accessed: Aug. 2023. [Online]. Available: https://www.smarthomesounds.co.uk/
[93] G. Terzopoulos and M. Satratzemi, "Voice assistants and smart speakers in everyday life and in education," *Informat. Educ.*, pp. 473–490, Sep. 2020.
[94] Selenium. *The Selenium Browser*. Accessed: Oct. 2023. [Online]. Available: https://www.selenium.dev
[95] Amazon. *Alexa Shopping List*. Accessed: Oct. 2023. [Online]. Available: https://alexa.amazon.com/spa/index.html#lists/namedLists
[96] *Speech Recognition Library*. Accessed: Sep. 2023. [Online]. Available: https://pypi.org/project/SpeechRecognition/
[97] K. Reese, T. Smith, J. Dutson, J. Armknecht, J. Cameron, and K. Seamons, "A usability study of five two-factor authentication methods," in *Proc. 15th Symp. Usable Privacy Secur.*, 2019, pp. 357–370.
[98] A. Ponticello, M. Fassl, and K. Krombholz, "Exploring authentication for security-sensitive tasks on smart home voice assistants," in *Proc. 17th Symp. Usable Privacy Secur.*, 2021, pp. 475–492.
[99] K. Tang, Y. Wang, H. Liu, Y. Sheng, X. Wang, and Z. Wei, "Design and implementation of push notification system based on the MQTT protocol," in *Proc. Int. Conf. Inf. Sci. Comput. Appl. (ISCA)*, 2013, pp. 116–119.
[100] A. Kumar and S. Johari, "Push notification as a business enhancement technique for e-commerce," in *Proc. 3rd Int. Conf. Image Inf. Process. (ICIIP)*, Dec. 2015, pp. 450–454.
[101] M. Tabassum, T. Kosinski, and H. R. Lipford, "'I don't own the data': End user perceptions of smart home device data practices and risks," in *Proc. 15th Symp. Usable Privacy Secur.*, Santa Clara, CA, USA, Aug. 2019, pp. 435–450. [Online]. Available: https://www.usenix.org/conference/soups2019/presentation/tabassum
[102] M. Tabassum and H. Lipford, "Exploring privacy implications of awareness and control mechanisms in smart home devices," *Proc. Privacy Enhancing Technol.*, vol. 1, pp. 571–588, Jan. 2023.

HASSAN A. SHAFEI received the B.S. degree in information systems from Jazan University, Jazan, Saudi Arabia, in 2012, and the master's degree in information technology from Lewis University, USA, in 2017. He is currently pursuing the Ph.D. degree with the Computer and Information Science Department, Temple University, USA. He is also an Instructor with Jazan University. His research interests include applied cryptography, security and privacy in IoT, and voice interface.

CHIU C. TAN (Member, IEEE) received the B.A. and B.S. degrees from The University of Texas at Austin, in 2004, and the Ph.D. degree from the College of William and Mary, in 2010. He is currently an Associate Professor with the Department of Computer and Information Sciences, Temple University. His research interests include cyber security, mobile AR/VR, camera-networks, smart health systems, and wireless network security (mainly 802.11, RFID, and sensor networks).