SoK: Context Sensing for Access Control in the Adversarial Home IoT

Weijia He, Valerie Zhao, Olivia Morkved, Sabeeka Siddiqui, Earlence Fernandes†, Josiah Hester‡, Blase Ur
University of Chicago, † University of Wisconsin–Madison, ‡ Northwestern University
{*hewj,vzhao,omorkved,sabeeka,blase*}*@uchicago.edu, † earlence@cs.wisc.edu, ‡ josiah@northwestern.edu*

***Abstract*—In smart homes, access-control policies increasingly depend on contexts, such as who is taking an action, whether there is an emergency, or whether an adult is nearby. The vast literature on context sensing could potentially be leveraged to support contextual access control, yet this literature mostly ignores attacks, adversaries, and privacy. In this paper, we reevaluate the literature on home context sensing through a security and privacy mindset. We first describe a novel threat model in smart homes focusing on the capabilities of non-technical adversaries. Replay, imitation, and shoulder-surfing attacks are much more likely in this model. We summarize contexts relevant to access control in homes, mapping them to existing sensors. We then systematize the sensing literature to construct a decision framework for home context sensing that considers security, privacy, and usability. Applying our framework, we find that current sensors do not fully mitigate likely threats in homes. Some sensors are susceptible to simple threats like physical denial-of-service attacks, making it easy to bypass policies relying on the absence of a characteristic. Many sensors collect more data than needed and are not effective for all groups of users or under all situations.**

***Index Terms*—access control, smart homes, context sensing, sensors, Internet of Things, IoT**

1. Introduction

Smart home technologies, including Samsung SmartThings [121], Amazon Echo [4], and Google Nest [45], are being widely deployed. While convenient, they introduce security and privacy risks [84], [103], [104], [118].

Access control is a key problem that requires new approaches in smart homes due to the inherent differences between devices and scenarios in the home IoT relative to traditional computing [55], [126], [161]. Traditionally, access to a resource is granted unconditionally to a user or a role encompassing a fixed set of users by typing a password. However, home IoT devices usually have a fixed location, while the users present vary over time. Thus, the set of users who ought to have access to a resource varies over time and may include guests, in-home workers, and others [29], [67], [161]. Users of shared devices may have complex social relationships, such as parent-child [124], parent-teenager [145], neighbors [15], or roommates [79]. Many home IoT devices also lack screens and keyboards, so interaction occurs through voice, gestures, or physical interactions like pressing buttons [43], [123]. Even for devices with screens (e.g., a smart TV), the content on the screen is often meant to be shared, putting users' password at risk. These new modalities require authentication mechanisms beyond passwords [27].

Crucially, He et al. [55] and Schuster et al. [126] found that desired access control in smart homes is frequently *contextual* (situational). Rather than granting unconditional access to a given user or a given role, authorization decisions may depend on the context. A context can be the user's location relative to the device, the history of the user's interactions with the device, or the state of the home [55]. An example policy is that a child can only use the smart TV when a parent is nearby [55]. Here, the system must verify two contexts: (i) a child is trying to use the TV and (ii) a parent is around. Enforcing contextual access control requires privacy-preserving and trustworthy context sensing. That is, a *sensor* (e.g., a motion sensor) must reliably detect some *context* (e.g., a room is unoccupied) while respecting users' privacy.

Prior work in the security and privacy community has already proposed ways to utilize contexts in access control [60], [126], but has not focused on how to detect contexts in the physical world in ways that are both trustworthy and privacy-preserving. A large amount of existing work on sensing and ubiquitous computing could be applied here, but it mostly ignores attacks, adversaries, and privacy. For example, work done on robust sensing often sacrifices privacy by adopting more invasive sensing methods [21] or denser sensor deployment [13], [94]. This is not realistic for an intimate setting like one's home. Some bodies of work also discover that errors are bound to occur in particular circumstances, but they regard these errors as rare or unintentional occurrences [61], [148], [171]. Adversaries can exploit this assumption.

In this SoK, we critically reevaluate the literature on context sensing in homes with a security and privacy mindset. Furthermore, we translate this literature to the problem of context sensing for access control, identifying sensor types that best match specific contexts within practical constraints. To do so, we first identified home contexts that are critical to access control from the small literature on contextual access control in smart homes. We then systematically searched the proceedings of the last decade of top conferences in sensing systems (SenSys, MobiSys, and MobiCom), ubiquitous computing (UbiComp/IMWUT), and human-computer interaction (CHI and UIST), identifying dozens of recent papers about sensors that can detect those contexts in smart homes. To capture well-known mature sensors, we also searched for commercially available sensors for smart homes and added classic papers on relevant sensors. This process left us with 94 pairs of contexts and sensors. Analyzing these papers while also revisiting key IoT papers from the secu-

rity, HCI, and usable security literatures, we constructed a decision framework that highlights each sensor's pros and cons for security, privacy, and usability when used to detect an access-control-relevant context in a smart home. Our work thus lays a foundation for secure, practical, and privacy-preserving context sensing in smart homes.

Our first contribution is a novel threat model broadening the adversaries that prior literature has considered for smart home sensing. Prior work has focused on how experts can exploit IoT systems through software vulnerabilities [3], [166], default passwords [5], replication of physical traits [89], and adversarial examples [17], [36], [130], [163]. While our model encompasses these threats, we focus on non-technical adversaries with legitimate access to a home, such as kids, roommates, guests, and workers, who usually have stronger motivations than remote strangers. Notably, most papers on context awareness and home sensing do not consider the adversarial mindset typical in the security community.

From our threat model, we make several observations. First, physical denial-of-service attacks are trivial against many sensors. Thus, in contextual access control, policies that allow access by default or rely on the absence (rather than presence) of a characteristic are easy to bypass. Second, non-technical users are highly capable of replay, imitation, and shoulder-surfing attacks. They can also impersonate someone by simply taking that person's phone. Identity cannot be reliably authenticated through possession of a phone or naive recognition of voices/faces.

Contextual access control in homes thus requires deploying sensors with key properties. The sensors, alone or in ensemble [13], must resist attacks from both technically literate outsiders and non-technical insiders. They must also minimize inadvertent data collection because sensors may be deployed in private areas of the home. Finally, household members must find the sensors acceptable.

Our second contribution is a decision framework for evaluating the degree to which a particular sensor possesses these key security, privacy, and usability properties. We further distinguish between attacks of different complexities, privacy considerations from various actors, and specific usability criteria. The latter includes ease of deployment, reusability of a sensor across contexts, and inclusiveness. This framework will be useful for individuals who design or deploy sensors in homes, including DIY users [16], manufacturers, and researchers in security and in sensing. We will refer to these individuals as *smart home designers*. This framework can help smart home designers navigate the vast array of sensing mechanisms described in the literature or available commercially. We envision the framework helping smart home owners to decide which sensor to use, manufacturers to design their products for facilitating contextual access control, and researchers to develop sensors that are more sensitive to security and privacy issues. The framework also outlines criteria to consider when designing a new sensor. In particular, our framework elucidates key trade-offs among the variety of sensors (e.g., motion sensors, microphones, thermal imaging) that can detect a given context (e.g., whether anyone is in a room).

For our third contribution, we apply our framework to highlight trade-offs in deploying sensors for access control in homes. Through a systematic review of the sensing literature, we identify *indicators* (e.g., characteristics, such as gait) and associated *sensors* (e.g., a pressure sensor mat for detecting gait) for sensing either *identity* (e.g., this is Jane) or *context* (e.g., this is an adult). Using our decision framework, we evaluate each sensor's key properties. We used our literature review to gauge sensors' robustness to attack, privacy properties (e.g., requirements for data storage), and deployability. With our framework, smart home designers can identify the sensors that support desired contexts for access control and recognize trade-offs in security, privacy, and usability. **To keep our framework and evaluations up-to-date, we have released them in a public GitHub repository.**[1] Researchers may publicly modify, expand, or dispute the table through pull requests and issues, facilitating open discussion between the sensing and security communities.

Applying our framework yields the following insights. First, we find that many current sensors, when used alone, do not adequately address potential threats from non-technical adversaries. They are especially vulnerable against rarely studied physical DoS attacks. Second, many sensors collect more data than needed. Contrary to currently deployed architectures, many sensors do not require cloud storage for data. Lastly, we found that many sensors are not inclusive based on age or disability, and some can be ineffective under certain environmental factors.

We first detail related work (Section 2) and describe our smart home model underpinning our framework (Section 3). This model identifies sensors that may operate in the smart home and would thus be in scope for our systematization. Our security and privacy evaluations of these sensors relied on our threat model of adversaries and attacks on context sensing, which we describe next (Section 4). Then, we propose our decision framework for selecting sensors for contextual access control in terms of security, privacy, and practicality (Section 5). We also describe our method for evaluating sensors through this framework (Section 6). Applying the framework, we make many observations about security and privacy in practical situations (Section 7). Finally, we discuss our framework's applications (Section 8) and conclude (Section 9).

2. Related Work

Privacy in the IoT: Prior research demonstrates gaps between users' perceptions of their privacy in smart homes and reality. Users have a limited understanding of privacy risks, and these opinions vary by context [10], [34], [77], [88], [161]. Researchers have explored the design space of usable privacy protection for the IoT [157], designed mechanisms for data transparency in homes [64], [65], [113], and made personalized privacy mechanisms that predict an individual's preference in a given situation [12], [34]. Building on the usable privacy literature, we capture privacy concerns about implementations of contextual access control, including sensing, data storage, and retention.

There is a growing body of research on developing privacy-preserving measures in ubiquitous sensing systems, such as obfuscation in audio sensing [42], cross-device tracking through ultrasound [91], bystander privacy in wearable cameras [31], [54], and so on. However,

1. https://github.com/UChicagoSUPERgroup/eurosp21

the public has no guarantee that the manufacturers of smart home devices will implement any of these countermeasures. Non-technical users are also unlikely to deploy such measures by themselves. Therefore, smart home designers should select a sensing method that minimizes overprivileged and inadvertent data collection before deployment, which is the focus of our paper.

IoT SoKs: A few prior papers survey specific aspects of the IoT. Fernandes et al. highlighted that access control and authentication are among the IoT's new intellectual challenges [38]. Considering software and networks, Alrawi et al. proposed methods for security evaluations of home IoT devices [3]. While they comprehensively explore digital attacks, we instead focus on physical attacks on sensors. Yan et al. examined analog sensor security, formalizing sensor circuits' security properties [156]. Their attackers are highly technical, whereas we focus on the non-technical adversaries that are common inside homes. On the network level, Yu et al. argued that context-aware enforcement is essential in the IoT [159]. Zhang et al. [166] compared academic and industry perspectives on IoT security. Their "environment mistrust" category can include physical attacks on sensing. They focus on technical attacks, such as signal jamming and voice synthesis. We expand their threat model to explore physical attacks on sensors by non-experts.

Access Control in Smart Homes: Researchers have studied users' mental models of access control for IoT devices [55], [70], [144] and data in homes [92], finding that current systems do not address the challenges unique to smart homes. Inspiring our work, Schuster et al. proposed protocols for enforcing contextual (a.k.a. "situational") access control [126]. They introduced Environmental Situation Oracles that answer queries about context. They did not, however, investigate physical sensing. We investigate the trustworthiness and usability of the physical sensing that necessarily underpins their oracles.

To improve the robustness of sensing, Birnbach et al. proposed ensemble methods that combine sensors to verify physical events in homes [13]. While they focused on techniques for sensor fusion, we provide a framework to help designers choose a set of sensors with complementary usability/privacy properties and abilities to resist attacks.

User-centered work focuses on the expression of policies, not enforcement. He et al. mapped potentially desirable policies [55] and Zeng et al. studied the user interface for expressing policies in multi-user homes [162], but gaps remain. Current designs rely on the integrity and availability of sensor data, which we show are not guaranteed. Privacy concerns can also make people unwilling to deploy certain sensors in homes [23].

While an expanding body of work proposes contextual access control, *no prior work has investigated how to realize such a system* using existing physical sensors.

Network Attacks: Remote adversaries can exploit weaknesses in protocols or software to attack home IoT devices without being physically near a home. These attacks, which are not unique to smart homes, include network-based denial of service (DoS) attacks [5] and exploiting weak or default credentials [72]. Sensor-based IoT devices are conceptual successors to wireless sensor networks and thus inherit most attacks against them [32], [62], [69]. They have been widely studied and are thus out of scope.

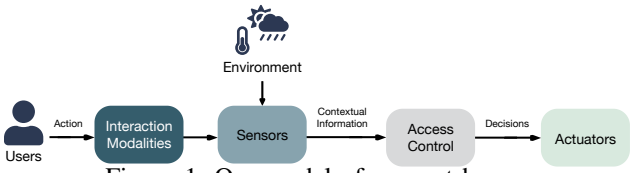

Figure 1: Our model of a smart home.

3. Our Model of a Smart Home

Context sensing and access control depend heavily on how a smart home works. Here, we abstract away implementation differences and discuss a model that applies to most smart homes. Current IoT devices support rich functionality, yet access control in the home has largely been limited to using smartphones as a proxy for identity.

Figure 1 depicts our basic model. A smart home consists of two types of Internet-connected devices: *actuators* that execute commands (e.g., lights), and *sensors* that measure their surroundings (e.g., motion sensors). Users control actuators through *interaction modalities* (e.g., smartphone, voice, physical buttons). The access-control policy uses contexts sensed via sensors to decide whether to authorize access.

Actuators can be controlled over the internet or a local network, enabling access control [126]. Traditional devices (e.g., non-IoT locks) are outside our model.

Users are people with remote or local access to devices, including family members, visitors, and workers.

Interaction Modalities describe how the user interacts with devices. Our model includes five modalities. The first four typically result in immediate changes, while the last covers automation that causes future changes.

1. Manual Interaction: A user can interact with devices manually, often by flipping switches or pressing buttons. Additional sensing is required to identify the user in such scenarios. A contextual access-control framework can inform a smart device whether to permit access.

2. Smartphones: Smartphone apps can control devices, sometimes via a home hub. Because users already authenticate to their phone, current IoT systems often rely on the possession of a phone as a proxy for identity.

3. Voice: Voice assistants let users control devices by speaking. Currently, they perform no authentication [154] or use speaker recognition that is easy to fool [132], [165].

4. Gestures: Currently uncommon in homes, gestures could be detected using ultrasonic or radio waves to recognize and authenticate movements as a source of input.

5. Automation: Smart home automation can link changes in context or other triggers to actions. They can be set with apps [66] or end-user programming [146]. Absent access control, automations may create loopholes [133], [150]. Imagine the automation: "If the lights turn off then play a movie." If a child may not play movies, yet may turn off lights, a crafty child could start a movie by turning off a light. While focused on contextual access control, our framework can also apply to automations triggered by a sensed context [146], such as when a room is warm [133], [164]. An attacker who tricks a sensor can cause chained automations toward a malicious goal [150], [151].

Contexts describe a particular state of the physical world. In a smart home, contexts describe situations, states of actuators, presence of specific people, and more. Examples include a security camera being activated, the temperature staying within some range, or a specific person sitting in the kitchen. Contextual access control relies on sensors to reconstruct these situations.

Sensors detect physical properties. Traditionally, they have been used primarily for smart home automation (e.g., motion triggers a light). However, recent research has identified the need for contextual access control in the smart home [55], [126], [162]. We envision that both existing and future sensors will underpin this paradigm.

Smart homes use phones or accounts as an imperfect proxy for identity. Context sensing has generally been used for automation, not contextual access control.

4. Our Threat Model

Sensor-based access control in homes requires robust sensing that protects user privacy. Prior IoT research has primarily focused on defending against remote attacks against IoT software [3], [159]. However, local attackers—regardless of technical background—can also pose a significant threat to the system by tricking physical sensors into detecting incorrect contexts or violating others' privacy. In fact, potential local attackers like family members, roommates, guests, and workers could have stronger motivations to bypass access control than unacquainted remote attackers. Our work examines local threats broadly and focuses on those posed by non-technical users with legitimate or illegitimate access to a home. Below, we taxonomize goals, attacks, and attackers. In light of the larger literature on context sensing, we revisit these attacks within our decision framework (Section 5).

4.1. The Attacker's Goals

One of our key insights is that non-technical attackers with modest and localized goals are a threat to contextual access control. Whereas remote attackers disrupt at scale, non-technical local attackers might only want to gain illegitimate access to some resource or spy on another individual. For example, a child may wish to watch TV without approval, a burglar may want to erase security camera footage after committing theft, or (as can be the case with intimate partner violence [41], [90]) an abusive member of the household may try to spy on members of their household by evading policies stopping security cameras from recording when people are home.

Local attackers might aim to bypass access control or compromise the privacy of others in the home.

Strategies for attacking sensors depend on the policy. A *default-deny* policy, which automatically denies access to unknown users, is not always advisable. For instance, prior work found users prefer default-deny policies for locks, but would rather permit unauthorized users to control smart lights than leave users in the dark [55].

Impersonation: Under a default-deny policy, a system only accepts authorized and authenticated users. An attacker must impersonate an authorized user or fabricate a valid token through imitation or replay attacks.

We find that these attacks often do not require technical knowledge (Section 7), especially in an intimate setting like a home where boundaries to privacy are reduced and private resources are easy to acquire. For example, many widely deployed facial-recognition systems lack depth or liveness detection. One can trick them by presenting a photo or video of an authorized user [87]. Photos of authorized users (e.g., a child's parents) are easy to find in a home, and videos can be taken in secret.

Similar issues arise for audio. People with access to a home can record authorized individuals speaking to voice interfaces. While authenticated speaker recognition is an active area of research [37], many widely deployed voice interfaces are vulnerable to simple replay attacks [132], [165] or even lack authentication entirely [154].[2] Off-the-shelf voice morphing compounds this problem [97].

Local attackers have extensive access to photos and audio, making basic face or speaker recognition systems vulnerable to replay and imitation attacks.

Current home IoT systems tend to rely on smartphones as a proxy for identity, capitalizing on their ubiquity. However, smartphones often run out of battery, and they do not offer the convenience of other interaction modalities (Section 3). This practice also falsely assumes that the user is always near their phone. For example, if the smart TV will turn on only if an adult's phone is in the room, a mischievous child can take their parent's phone while the parent is sleeping. Furthermore, smartphone authentication is still not fool-proof as it is often knowledge-based (e.g., PINs). It is often easy for others in the home to bypass this authentication through shoulder-surfing.

Existing practices of using phones (potentially with authentication) as a proxy for identity in shared spaces can be risky in terms of both security and usability.

Invisibility: Contextual access-control policies can also allow access by default. One example would be using the smart stove. Whereas visitors or babysitters may be allowed to use the stove, a child should not use it for safety reasons. A natural policy that follows is "anyone except a child can turn on the stove." When these *default-allow* policies depend on *not* sensing a characteristic or situation, e.g. "record security video of the bedroom when *no one is home*), an attacker needs nothing more than to make the characteristic or situation "invisible." They can do this by changing or blocking the sensor's field of view.

We will refer to such attacks, where the local attacker prevents the sensor from physically detecting a context, as *physical denial of service (DoS)*. This can entail blocking a motion sensor with paper or overloading a microphone with loud noise (including outside the human hearing range [1], [163]). Sensors must detect whether they are receiving accurate and fresh input.

Default-allow policies, which rely on *not* detecting a given situation, can be defeated by blocking sensors.

2. In our informal testing, Google Home's speaker recognition only seemed to verify the person who said "OK, Google." It accepted further commands spoken by someone else, making replay attacks trivial.

Dimension	Type	Capabilities	Examples
Access	Indoors	Physical access to indoor & outdoor devices/sensors Rich observation opportunities Full knowledge of sensor models & locations Knowledge of access-control policies & automations	Family member, babysitter
	Outdoors	Physical access only to outdoor devices/sensors Limited observation opportunities Opportunistic attacks that reach more victims	Neighbor, prospective burglar
Expertise	Expert	Sophisticated network and imitation attacks Ability to craft black-box adversarial examples Unsophisticated replay/imitation attacks, block sensor	IT professional, hacker
	Non-expert	Unsophisticated replay/imitation attacks, block sensor	Child, domestic worker
Resemblance	Similar	Spoofing (through imitation) Higher possibility of inadvertent false positives	Sibling, one who looks similar

TABLE 1: Local attackers can be characterized along the dimensions above, impacting attack capabilities.

4.2. Attacks

Based on these attacker goals, we surveyed top security and sensing conferences to identify likely attacks. We clustered prior work based on attack method, resulting in three major types of attacks: 1) replay and spoofing attacks; 2) adversarial examples; 3) sensor hardware attacks. Note that replay and spoofing attacks differ in practicality despite often appearing together in the literature. We did not find mentions of physical DoS attacks in our literature survey, but include them because they are a clear threat to access control. Below, we define these attacks.

Replay Attack: The attacker collects a credential and feeds it back to a sensor. For example, the attacker can play a voice recording, show a photo of a face, or make a gummy mold of a specific fingerprint [89]. Our focus in this SoK is on replaying the physical signal itself, although network traffic can sometimes also be replayed.

Spoofing: The attacker forges an approximate credential or situation they have not necessarily captured. Smoke can spoof a fire, and energetic pet cats can spoof occupancy.

Physical Denial of Service (DoS): Jamming, blocking, or moving a sensor can prevent accurate sensing. It is important to note that the sensor detecting the *absence* of a characteristic or situation is different from *not* detecting it. For instance, when trying to sense whether a room is empty, a camera blocked by a piece of paper will not detect any people. This differs from a camera affirmatively seeing a room without people. These attacks are often easy to deploy, but have not yet received much attention.

Adversarial Examples: Against ML-based sensing methods, the attacker can poison the training data or add carefully crafted noise to inputs [125], [130].

Sensor Hardware Attacks: The attacker leverages the physical principle behind the hardware to deceive the sensor, such as with signal injection attacks [73], [163].

Inadvertent False Positives: This is not quite an attack, but a sensor incorrectly detecting an identity or situation can still compromise access control.

4.3. Physical Sensors' Potential Attackers

To understand each attack's feasibility, we characterize the attacker's capabilities. Table 1 provides a summary. Our threat model concerns attackers who *violate* access-control policies. We thus ignore adversaries who *create* unreasonable policies, such as domestic abusers attempting to spy on their family. Defending against those adversaries requires countermeasures beyond access control.

Access: An attacker with access to the home would be well-positioned for physical attacks against sensors. They can observe authentication processes in the home, potentially repeatedly, to record information for replay or imitation attacks. For example, a roommate might encounter multiple instances of the user speaking to a voice assistant. They thus have multiple opportunities to record the user's voice for tricking speaker-recognition algorithms. By having access to the home, attackers can also infer access-control policies, automations, and sensor locations or types from their observations. Legitimate access can be permanent, such as for residents, or temporary, such as for visitors and domestic workers. Illegitimate access occurs when people enter the home without permission.

It is also possible for attackers to access sensors outside the home [15], [82] or make inferences using partial information (e.g., from sensors visible through windows). Some individuals who might rely on these methods include neighbors and prospective burglars. We note that modeling the attack surface cannot rely on a simple indoor versus outdoor dichotomy. For example, one can control a voice assistant through an open window.

Expertise: Attackers with technical expertise, such as infosec professionals, are capable of sophisticated attacks. Some attacks against ML-based sensor systems are of this nature. They can involve carefully crafted eyeglasses [130], stickers [36], or audio [1], [163]. Experts can also target sensors' physical principles, such as applying acoustic interference to accelerometers [141]. Finally, network- and software-based attacks are also possible.

On the other hand, nontechnical attackers can carry out replay or imitation attacks that only require observations (e.g., spoken passwords) or commodity recording equipment (e.g., a smartphone). They can also disable sensors by blocking, repositioning, or unplugging them.

Resemblance: Biometric sensors may confuse individuals of similar physical traits. Biological family members often share physical resemblances and have easy access to sensors because they often live together or visit each other. Real-world examples include one man who tricked a voice-recognition system by imitating his twin's voice [132]. Identical twins can also fool facial recognition [143]. It may also be possible for unrelated people with physical resemblances to trick the sensors.

Our threat model highlights two key ideas missing from prior work. First, most work focuses on threats from attackers with extensive resources and expertise. We show that non-experts with access to the home are capable of replay and spoofing attacks against sensors that support contextual access control. Second, blocking sensors can allow attackers to evade some access-control policies. This method of attack has not yet been studied extensively.

Contextual access control must consider that non-experts with access to a home can attack sensors.

5. Decision Framework for Context Sensing

Individuals designing or deploying home sensors need a framework that helps them navigate the trade-offs between sensors' security, privacy, and usability properties in conjunction with the users' needs and the space itself [23]. These individuals, whom we term *smart home designers*, will benefit from the framework in different ways:

- Do-it-yourself smart home owners can learn security and privacy implications of selecting certain sensors.
- Sensor manufacturers can holistically evaluate their current sensors' trade-offs and identify additional contexts that need new sensors to be developed.
- Security and sensing researchers can identify security and privacy gaps that guide their future research.

For example, a smart home owner might wish to know when anyone is at home. Consulting our framework reveals that cameras are suitable for this, but are not privacy-preserving. Meanwhile, pressure sensors on the floor would be privacy-preserving, but are impractical and expensive to install. The user can now determine whether to prioritize occupancy detection at the cost of privacy.

Here, we first explore the life cycles of adopting a sensing technique. Then, for each stage of the life cycle, we further define the main security, privacy, and usability criteria that smart home designers must consider in choosing sensors, which we collectively consider our framework. We constructed this framework by critically analyzing the 94 pairs of sensors and contexts we identified through our systematic review of the sensing literature (see Section 6.2) relative to the security and usable security literatures concerning the home IoT (see Section 2). We also considered broader security principles to fill in potential gaps in this framework.[3]

5.1. Life cycles

Adopting a new sensing technology in one's home is a long-term and ongoing process. To avoid missing crucial challenges during the process, we first define different stages of the adoption process, as depicted in Figure 2.

Acquiring the required hardware: A user might need to buy new sensors, which is a financial and time investment.

Deploying the hardware: After acquiring the hardware, users need to install it in their homes. When needed, users might also re-deploy hardware, such as to reposition it.

3. The team that constructed the framework included multiple students and three faculty members. Two of the faculty members focus on security and privacy research, but also have experience with machine learning research. The other faculty member conducts sensing research.

Registration (optional): Sometimes the hardware may require the user to register themselves first, which is especially common for sensors pertaining to an identity.

(Re)training / Maintenance (optional): Before usage, machine learning-based sensing methods commonly require the user to train the model about the context in its unique environment. Retraining may also be required in the future to adapt to users and a sensor's environment changing over time. Maintenance, such as battery replacement and routine check-ups, may also be required.

Usage: After training, the sensor is ready for use. We expect the sensing technique to operate until the user stops using the sensor. To identify possible issues in this stage, we must abstract how the sensing technique works.

Sensing detects environmental events, such as temperature changes, movement in the background, and sound. We term these *indicators*, which could be mapped to a context. For example, if a sensor detects movement of a heat source, it is likely to be someone moving nearby.

To detect the indicator, the sensor needs a signal sent or radiated from the source. Depending on how far the signal can be transmitted, the sensor may require direct contact, near-field communication, or far-field communication. We term this process *signal transmission*.

Once the sensing hardware receives the signal, it first needs to process the analog signal, such as using amplification and noise filtering. The analog signal can then be converted into a digital signal for further processing. The *sensing hardware* stage represents the above process.

Finally, the digital signal, or the raw sensor data, is sent to a processor or the cloud for further computation. Depending on what the sensing method is designed for, different *data analysis* methods may apply here. For example, facial recognition and gait recognition may both rely on cameras, but the data analysis would differ. Once the sensed data analysis is complete, the algorithm *outputs* whether the context it aims to detect is active.

End of Life: The user may eventually decide to uninstall the home sensor. In this stage, the sensor may be directly thrown away, given to others, or sent back to the manufacturer for upgrading or replacing. The hardware is not guaranteed to be properly destroyed. Thus, information leakage after disposal is possible. We treat the uninstallation process as two parts: removing all data (e.g., factory reset) and physically removing the sensor from the home.

5.2. Security

We consider two ways in which a sensor may be attacked. One is through *inadvertent failures*. An attacker may bypass an error-prone sensor through brute force. The other is through *intentional attacks*. These attacks are described in detail in Section 4. Figure 2 also indicates at which stage these attacks might occur.

We do not consider attacks before the usage stage. The set-up stage occurs only once and the victim is often present, increasing the difficulty of attacking the sensor itself. Therefore, during the set-up stage, it is more likely for the attacker to perform network attacks (e.g., sniffing, person-in-the-middle), which are out of this paper's scope.

In Tables 2-3, a red "!" signifies that a sensor is easily susceptible to a given attack. A yellow "?" signifies that it is not very susceptible to the attack. If no symbol is

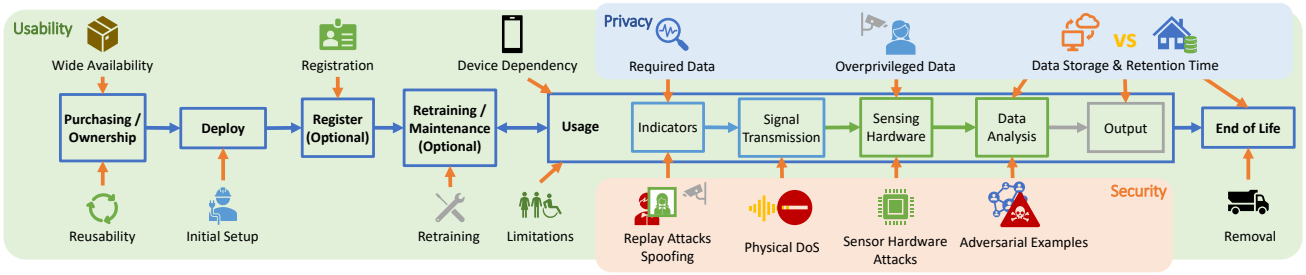

Figure 2: Different issues emerge in difference stages of using sensors in home.

shown in the table, the attack is implausible against the sensor (e.g., replay attacks against smoke detectors).

5.3. Privacy

Sensors collect data to operate, but excessive collection of sensitive data causes privacy concerns. Furthermore, certain contexts require intensive computation on data that is collected over long periods of time. To identify potential privacy threats during the usage stage, we review each stage carefully to identify general threats. We assume that the sensing software is secure and do not consider privacy threats before the usage stage. Our framework considers the following aspects:

Required Data: Data that must be collected for the sensor to function. Depending on which *indicator* the sensor detects, different types of data are collected, with various privacy implications.

Overprivileged Data: Depending on which sensor the designer decides to use, superfluous data might be collected inadvertently. For example, a microphone for occupancy detection also records conversations. In the "overprivileged data" column of Tables 2-3, *poor* means the sensor collects unnecessary and sensitive information, *acceptable* means it collects unnecessary data that is not sensitive, and *good* means it does not collect superfluous data.

Data Storage: Data must be analyzed and stored in the cloud if the device lacks the computational power or storage space for local processing. For other sensors, however, data can be stored on the device containing the sensor or on an in-home hub. Nonetheless, companies often upload data to the cloud even when unnecessary [28]. There is no guarantee that the uploaded data will be used ethically [56], which can deter users from deploying some sensors in homes [23]. We consider whether each sensor's data *must* be stored on the *cloud*, or whether *local* storage supports the needed functionality. We leave out of scope the question of whether a company will choose to upload data to the cloud even when it could be retained locally.

Retention Time: Some sensors require longitudinal data (e.g., for training a model). Companies may again decide to store all data indefinitely even when not strictly necessary. *Transient* storage means sensed data can be immediately discarded, while *persistent* means it must be retained until the user factory resets the device. Similar to *data storage*, companies may retain users' data for as long as they want, even if the user factory resets their device and deletes their account. To focus on the requirement for enabling the sensing technology, we only consider how long the data must be available for the functionality.

5.4. Usability

To assess a sensor's usability for a non-technical end user, we consider the following criteria, which we compiled based on the stages identified in Figure 2.

Wide Availability: Users are more likely to adopt sensors that they can easily acquire. For example, one can sense occupancy with motion sensors or ultrasonic sensors, but users and designers may prefer the former because of their cheap cost and ubiquity. Nonetheless, more expensive sensors (e.g., cameras) may also be widely available if they fulfill multiple use cases. This may benefit users because sensors that fulfill multiple use cases may obviate the purchase of additional sensors.

Initial Set-up: How difficult is it for a non-technical user to set up the hardware during the deployment stage? *Good* means little to no effort is required, such as plug-and-play installation. *Poor* requires substantial effort from the user, such as renovating their current home for installation (e.g., painting the wall, changing the floor). Anything between *good* and *poor* was deemed *acceptable*.

Registration: How much effort does it take to register a user, or how long does it take to collect enough data to train the model? *Good* means no registration or training is needed. *Acceptable* encompasses two situations. In the first situation, the sensing method requires straightforward registration or data collection, meaning registration should not take over 10 minutes. This includes most commercial products, such as Touch ID or Face ID. In the second situation, data collection needs more time to finish, but does not require user attention. For example, a system from Hsu et al. [61] required the user to wear an accelerometer for days as ground truth for identifying the user from their RF reflection. While this process takes days, no attention is required, earning it an *acceptable* rating. *Poor* takes significant effort from users, usually exceeding 10 minutes in duration while requiring constant attention the entire time. For example, Qian et al.'s system [111] requires the user to walk for four minutes each at three different paces.

Retraining / Maintenance: How often is model retraining or hardware maintenance required? *Good* requires none. *Acceptable* requires occasional retraining or maintenance less than once a month (e.g., changing batteries every few months). *Poor* requires retraining or maintenance at least once a month. When evaluating biometric sensors, we assume an adult user with stable features.

Reusability: Some sensors can detect multiple contexts. For example, cameras can detect age, room occupancy, or an identity. *Good* means many contexts can be sensed, as with cameras. *Acceptable* means a few contexts can

be sensed, as with radar sensors. *Poor* means the sensor detects only one context, as with fingerprint sensors.
Device Dependency: Some methods require users to carry a device (e.g., a phone) during usage. *Good* means no such device is required. *Poor* means that it is required.
Limitations: We consider whether the sensor is effective for all groups of users and under all situations. We focus on age, potential disabilities, and environmental factors (e.g., lighting conditions, GPS reception underground).
Removal: When a user decides to stop using a sensor, the sensor will be removed from the home. As removal is the inverse of the initial setup, we decide to combine them with the initial setup in Tables 2-3.

5.5. Example

We illustrate the use of this framework by describing two examples. Both examples are sensors that one might use to detect robbery, which is relevant to when access is granted based on whether there is an emergency in the home. They are also listed in Table 2.

Some commercial products, such as the Netatmo Camera [102], alert the user when unrecognized individuals enter the house. As one would expect, cameras and facial recognition algorithms have poor security and privacy qualities, but great usability. They are easily susceptible to replay attacks and adversarial examples. They are also susceptible to physical DoS if the attacker simply blocks the field of vision with an object. Sensor hardware attacks and spoofing are likely impossible for the adversaries we consider. The video stream will capture more information than needed to determine the occurrence of a robbery. Processing the video stream requires long-term cloud storage. Lastly, cameras are ubiquitous and easy to use, although registering users and retraining the facial recognition algorithm to accurately recognize users require some effort.

Glassbreak sensors, like Honeywell's [58], can also detect robbery by monitoring for audio frequencies of glass breaking. These sensors are susceptible to replay attacks, physical DoS, sensor hardware attacks, and spoofing. Machine learning is not necessary, so adversarial examples are not a concern. They capture basic audio frequencies that encode more information than necessary, but this information is simple enough to be stored locally for a short amount of time. They are easy to acquire and use, but they only fulfill the unique purpose of detecting glass breaking. A user looking to sense multiple contexts cannot rely on glassbreak sensors for other contexts.

6. Methodology

Both to understand the potential of applying our decision framework in realistic situations and to illustrate how to use it, we applied the framework to sensors that would support commonly desired contextual access control policies in smart homes. Applying the framework requires: (1) a set of desirable contexts for access control policies; (2) sets of sensors that can detect those contexts; and (3) evaluations of the security, privacy, and usability of detecting those contexts with those sensors. This section details our method for applying the framework and analyzing each aspect to create Tables 2–3.

6.1. Desirable contexts

Existing work on context sensing does not fully list the desirable contexts for contextual access control in homes. For example, some work focuses on non-security domains, such as sensing contexts for healthcare [95], activity recognition [76], [169], or indoor tracking [68], [107], [112], [155]. Other work focuses on device-level contexts (i.e., device states) [18], [66], [139], [159], but does not consider contextual access control.

To overcome these challenges, we first identified a list of contexts mentioned in the most closely related work on contextual access control in homes [55], [126], [162]. We then analyzed the user study data from He et al. [55]. We manually clustered participant responses through open coding. We added to our list contexts mentioned at least five times or that are related to identity (thus naturally relating to access control). Tables 2–3 list the final set of desirable contexts in the leftmost column. The "user" in the leftmost column refers to the initiator of the action who uses a device that is owned by the "owner."

6.2. Sensing Mechanisms

Extensive prior work proposes technologies to sense identity or contexts in physical spaces. It is hard for a smart home designer to navigate this work and determine the appropriate sensor based on its security, privacy, and usability trade-offs. For example, to track a person's location in the home, researchers have used cameras [155], CSI (Channel State Information) from WiFi signals [112], visible light channels [81], and more. Direct mappings between contexts and precise sensors are not straightforward. Generally, a physical *sensor* is used to sense some characteristic (which we term an *indicator*) that relates to that context. For example, if age is the relevant context, one might use a person's gait, voice, or facial characteristics as physical indicators of age. These indicators can be sensed with cameras, microphones, and more.

For each context, we identified potential indicators and associated sensors by surveying the sensing literature, searching for relevant industry products, and asking experts from the sensing community for methods they had encountered in their field. Our final set of sensors (see Tables 2–3) includes both research prototypes and mature products. The *example* column of Tables 2–3 lists the examples of research prototypes or commercial products we consider for each type of sensor.

To find and evaluate research prototypes, we systematically reviewed the last ten years of proceedings of top conferences in sensing systems (SenSys, MobiSys, and MobiCom), ubiquitous computing (UbiComp/IMWUT), and human-computer interaction (CHI and UIST) in the ACM Digital Library. We first filtered the search results based on keywords ("sensing" in the abstract and "home" in the paper), which yielded 716 papers. We then manually inspected each paper to determine its relevance. We used the paper's title to determine potential relevance, which led to 127 papers remaining. We then read each of these papers to determine its actual relevance. We further excluded papers if (i) they were not related to sensing in homes, but rather applications like VR/AR, smart cities, or health; (ii) they did not focus on sensing a specific

context, but rather on refining sensing techniques through improved processing algorithms or machine learning techniques; or (iii) we could not directly map the paper to any of the desirable contexts we identified. The final 36 papers are listed in Table 2, and we extracted the indicators of the contexts from the corresponding papers. If we did not find prototypes in this body of literature for an indicator, we looked to related top-tier conferences, such as CVPR.

To augment this initial list with more mature and commercially viable methods, we first consulted experts in the sensing community to identify classic papers for types of sensors that are now commonly used. To cover methods used in commercial products, we then searched for sensors of each indicator (as collected from research papers above) on Amazon. If we had not found any indicators at that point for a context, we searched for sensors related to that context and then included the indicators they used. This process led to our final set of 94 pairs of a context that is desirable to sense for access control in the home and a type of sensor (research prototype or commercial product) that identifies that context.

The steps described above survey, but do not systematize, this work. For systematization, we applied our framework to analyze the security, privacy, and usability of using that sensor to detect that context. To understand how the sensing method worked, we read the relevant research papers for prototypes and any user manuals, technical specifications, and white papers we could find for commercial products. We list the detailed criteria we use for this systematization below and in Section 5.

6.3. Security

Attacks, listed in Section 4, target particular types of sensors. To perform replay attacks, one must be able to record and then play back the relevant data, a situation that mostly applies to microphones and cameras. Attacks on sensor hardware target sensors' physical properties and are thus relevant to microphones, MEMS sensors, and more. We used past literature to decide whether the type of sensor used by the sensing method is vulnerable or not.

Some attacks (e.g., physical DoS attacks) are less studied and some sensors (e.g., motion sensors) are less often targeted. In these cases, we studied the sensor's basic principles from papers, product manuals, and white papers, and we discussed among our team whether it might be susceptible to each attack. For example, passive infrared (PIR) motion sensors detect motion based on changes in their view in infrared. Infrared radiation struggles to travel through paper, glass, and thermal blankets, which makes occlusion possible. We acknowledge that some products may adopt anti-tampering techniques not specified in the manual or technical specifications. Our judgments reflect contemplation, rather than lab testing. Tables 2-3 thus outline expected and potential attacks.

6.4. Privacy

We evaluated sensors' privacy implications as follows. We identified the data required by each sensor based on its description in its paper or manual. Examples include audio for microphones, air for smoke detectors, and phone packets for CUPID [129], a WiFi-based indoor localization system. We then identified overprivileged data collection by subtracting the information needed to determine the context from what could reasonably be inferred from the required data. We used the guideline in Section 5 to label overprivileged data in Tables 2-3. For example, Touch ID [9] requires fingerprints. This might suggest overprivilege because a fingerprint is personally identifiable. However, since it is used to detect the user's identity, we do not consider its data collection overprivileged.

Next, we determined the data storage location and retention time required for reasonable performance. For storage location, we examined the algorithms needed to process the data for the sensor. If the sensor required a large amount of longitudinal data or algorithms that could not be computed locally (such as Gaussian models), we labeled the sensor as requiring *cloud* storage. Otherwise, we labeled it as *local*. For example, we consider local storage sufficient for sensors that use SVM classifiers and require only highly limited longitudinal data. If data did not have to be stored for more than one access, we labeled it *transient*. If any data did, then we labeled it *persistent*. For example, smoke detectors have transient data retention because they do not need to store historical air data to detect future smoke. In contrast, fingerprint readers that verify identity do need to store representations (templates) of the fingerprint to perform future matching algorithms.

6.5. Limitations

Due to a lack of access to many of the products and prototypes in our evaluation, the ratings we give are based on team discussion and contemplation. To the best of our ability, we tried to make the criteria as concrete as possible and to review papers and specifications with care. However, some cells in Tables 2–3 could be subjective and debated by researchers with different assumptions and access to different information. As such, we intend Tables 2–3 to reflect an initial attempt of applying our framework and distilling the pros and cons of each sensor in each context. We intend these tables as a living document that evolves with community effort and robust online debate, as we discuss further in Section 8.

7. Insights From Applying the Framework

We present key findings from applying our framework (Section 5) to sensors that support commonly desired contextual access-control policies in smart homes. Tables 2–3 summarize each sensor's pros and cons in security, privacy, and usability regarding detecting a given context.

7.1. Robustness to Attacks

Most sensors are vulnerable to physical DoS. Of the 94 context-sensor pairs evaluated, 64 (68.1%) are vulnerable to physical DoS attacks. Vision-, audio-, heat-, and EM-wave-based sensors (radar, WiFi, radio) can easily be blocked or jammed even by those with no technical background. Vision and heat-based sensors' line of sight can be blocked. Playing loud music floods audio sensors. Energy-absorbent materials can be placed near transmitters (e.g., black material near light-based sensors). Through these

Contexts	Indicators	Sensor	Example	Security: Error	Security: Replay Attacks	Security: Adversarial Examples	Security: Physical DoS	Security: Sensor Hardware Attacks	Security: Spoofing	Privacy: Required Data	Privacy: Overprivileged Data	Privacy: Data Storage	Privacy: Retention Time	Usability: Wide Availability	Usability: Initial Set-up / Removal	Usability: Registration	Usability: Retraining	Usability: Reusability	Usability: Device Dependency	Usability: Limitations
User's identity	Voice	Microphone, inertial sensors	[37]	0.1%					!	A,Bm										
		Microphone-only	[75]	5-6%						A, C, M										
			[46]†	–	!	!	!	!	!	A										
	Breathing patterns	Microphone	[19]	0.4-2%	?	!				A										
	Facial features	Camera	[102]†	Variable	!	!	!			V										
		Depth camera	[8]†	<0.001%		!	!			P'										
		Infrared (IR) camera	[93]†	<0.001%			!			P'										
		Camera, inertial, light sensors	[20]	4.7%			!			V, C, E										
	Eye features	Iris scanner	[119]†	–	!		!			P'					–					
	Fingerprint	Fingerprint sensor	[9]†	0.002%	?					F										
		Microphone	[116]	2-16%			!	!		A					–					
	Body shape	Radar (RF) sensor	[68]	10-21%			!			B										
	Bioimpedance	Bioimpedance sensor	[26]	2%						El										
			[122]	11-21%						El					–					
	Cardiac motion	Radar sensor	[83]	1.39%			!			Bm										
		Camera	[85]	1.4-4.5%			!	!		Bm										
	Hand gestures	IMU sensors	[115]	10-36.2%			!	!		M										
	Gait properties	Vibration sensor	[109]	10%			!			G					–					
		Load cells	[107]	7%	?					G										
		Pressure sensors	[111]	7.7%	?					G										
		Camera	[149]	6.25%	?		!		!	V										
		Microphone, WiFi TX & RX	[21]	8%-28%			!			C, A										
		Photointerrupters	[160]	1%			!			G										
Owner / guest	Identity	*Similar to "Identity" above*								*Similar to "Identity" above*										
User's age	Voice	Microphone	[128]		!	!	!	!	!	A										
	Facial features	Camera	[105]	6.01 - 6.08 yr.	!	!	!			P										
			[170]	4.83 - 6.28 yr.	!	!	!			P										
			[108]	2.514 - 3.086 yr.	!	!	!			P										
			[30]	22.24 - 9.07%	?	!	!			V										
Emergency in the home	Fire	Smoke detector	[47]†	Variable			!		!	E										
			[39]†	Variable			!		!	E										
		IR Camera	[140]	Variable			!			V'										
		IR/UV detector	[131]†	Variable			!			E										
	Toxic gas	Combustible gas detector	[137]†	Variable						E										
		Carbon monoxide detector	[44]†	Variable						E										
	Robbery	Camera	[102]†	Low	!	!	!			V										
		Glassbreak sensor	[58]†	Variable	!		!	!	!	A										
User in same *house* as the device	Presence of tags	Bluetooth Low Energy (BLE) signal sensor	[7]†	–	!		?		!	L'										
		RF/Ultrasonic sensors	[110]	–			!		!	L'										
		RFID	[127]	–			!			L'										
		WiFi TX & RX	[75]	10%			!		!	A, C, M										
	Movement	WiFi TX & RX	[148]	0.5m - 1.1m			!			C										
			[106]	1.84m			!			C, Fp					–					
			[152]	4%			!			C										
	Trajectory	Inertial sensors in phones	[80]	1.5 - 2m	!			!		G, M										
User in same *room* as the device	Presence of tags	BLE signal sensor	[7]†	–	!		?		!	L'										
		BLE, IMU sensors	[24]	2.42 - 14.72%			!	!		M, T, O										
		RF Techniques	[110]	–			!		!	L'										
			[168]	0.06%			!			D										
		IR tags	[153]	Variable	!		!		!	L'										
		Ultrasound TX & RX	[78]	0.1m	!					L'										
			[2]	3cm	!					L'										
		Capacitive NFC	[52]	–						L'										
		Visible Light Channel	[167]	5.9cm			!			L'										
	Movement	WiFi TX & RX	[148]	0.5m - 1.1m			!			C										
			[106]	1.84m			!			C, Fp					–					
			[152]	4%			!			C										
		Motion sensor	[127]	0.5m - 1.1m			!			M										
			[99]†	1.84m			!	!		M										
	EMI	Voltage sampling	[53]	6%						L'										
		Passive magneto-inductive sensors	[147]	6-17.4%						L'										
	RF reflection	RF sensor	[61]	81%			!			L'										
	Electric potential	Electrical potential sensors	[51]	0.16m			!			El										
	Location semantic	WiFi, microphone, IMU sensors, Barometer	[168]	0.627-0.778 (F-measure)			!			L'										
	Hand gestures	IMU sensor	[115]	83-91% (supervised)				!		L'										
	Water pressure	Pressure sensor	[138]	17.31% - 29.89%						L'										
Owner away or not	Location	GPS	[101]†	Variable			!		!	L										
Adult nearby	Age	*Similar to "Age" above*								*Similar to "Age" above*										

Note: In the "Example" column, † denotes commercial sensors or systems.

TABLE 2: An example application of our framework to sensors and contexts identified in our review of the literature and current sensing products. 36 of these sensors come from the academic literature, while the rest are commercial products, denoted with a † in the "Example" column. We mapped the sensors to contexts they are able to detect for the purpose of an access-control policy allowing or denying usage. The "Error" column contains reported values from the cited example sensors. Other columns reflect our best judgment, which was informed by the cited works when related information was reported. !/?/(blank) = Easy/Hard/Impossible, 👍/👍/👎 = Good/Adequate/Poor, 🏠/☁ = Local/Cloud, ○/● = Transient/Persistent data retention, – = Not found. For *Required Data*, **A** = Audio, **B** = Body shape, **Bm** = Body movement, **C** = CSI, **E** = Environment, **El** = Electrical properties of body, **F** = Fingerprint, **G** = Gait, **L/L'** = Geo/Indoor location, **M** = Movement, **P/P'** = Photo/Infrared photo, **D** = Device info, **V/V'** = Video/Infrared video, **T** = Temperature, **O** = Orientation, **Fp** = Floor plan. The rows of this table continue in Table 3.

Contexts	Indicators	Sensor	Example	Security: Error	Replay Attacks	Adversarial Examples	Physical DoS	Sensor Hardware Attacks	Spoofing	Privacy: Required Data	Overprivileged Data	Data Storage	Retention Time	Usability: Wide Availability	Initial Set-up / Removal	Registration	Retraining	Reusability	Device Dependency	Limitations
No one nearby	WiFi signals	WiFi TX & RX	[152]	96% (TPR)			!			C										
			[11]	Low	!				!	D										
	Presence	RF sensor	[68]	High	?		!			B										
		Camera with IR LEDs	[48]†	Variable			!			V'										
			[14]	Variable						V'										
		Load cells	[107]	7%						G										
		Pressure sensors	[111]	7.7%						G										
		Ultrasonic sensors	[57]	10%			!			B										
	Movement	Motion sensor	[120]†	Variable			!			M										
			[100]†	Variable			!	!		M										
			[59]†	Variable			!	!		M										
	Footsteps	Microphones		Variable	?			!	!	A										
		Similar to "Gait" above								*Similar to "Gait" above*										
	CO_2	Nondispersive Infrared (NDIR) CO_2 sensors	[25]†	Variable						E										
	Body heat	Infrared sensors	[50]†	Variable			!			V'										
People asleep nearby	Movement	Inertial sensors	[6]†	Variable				!		M										
			[40]†	Variable				!		M										
		Similar to "Motion sensors" above								*Similar to "Motion sensors" above*										
		Radar sensor	[114]	89.6% (recall)			!			M										
People present in same *house* as the user	Location	GPS	[101]†	Variable			!		!	L										
	Movement	Static electrical field	[96]	1.88%			!			E							–			
		RF sensors	[171]	Low			!			M										
	Presence of tags	RF/Ultrasonic sensors	[110]	–			!		!	L'										
		BLE signal sensor	[7]†	–	!		?		!	L'										
People present in same *room* as the user	WiFi signals	WiFi TX & RX	[135]	Variable			!			L'							–			
			[129]	1.8m	!		!			C										
	RF reflection	RF/Ultrasonic sensors	[61]	19%			!			L'										
	Sound (chat)	RF/Ultrasonic sensors	[75]	26%			!	!		L'										
	Doorway activity	RF/Ultrasonic sensors	[57]	10%			!			B										
	BLE signals	BLE signal sensor	[7]†	–	!		?		!	L'										

TABLE 3: A continuation of the rows of Table 2, which is an example application of our framework to the sensors and their associated target contexts. The abbreviations used are the same as defined in Table 2's caption.

means of hindering sensor operation, attackers can become invisible to systems with default-allow policies.

Physical DoS is hard to detect because the symptoms can be similar to normal activities. This is very different from network DoS attacks. Monitoring may alleviate the issue, but home occupants are unlikely to perform constant monitoring. A blocked sensor may not be noticed until the attacker has already achieved their goal.

Sensor redundancy can mitigate physical DoS attacks. For example, a room could have a motion sensor, a pressure sensor in the floor, and a microphone to detect whether the room is occupied or not. If access is granted when the room is unoccupied, an attacker wanting access would need to accomplish the difficult task of occluding all three sensors around the same time. By cross-checking the sensors' data streams with each other [13], the system could verify whether the room is unoccupied and determine whether a sensor has been compromised.

Careful policy design is another defense against physical DoS attacks. A system's default policy—whether to *allow* or *deny* access when a condition is met—can impact attack success. For example, a user might specify "my child should not have access to the TV." With a default-allow policy, TV access will be granted unless a child is detected, yet the child can block a sensor to avoid detection. With a default-deny policy, the child cannot rely on physical DoS.

The optimal default policy may vary based on the device or operation. Users may prefer default-allow rules for controlling lights because falsely allowing operation is typically of little consequence, but falsely denying operation causes inconvenience [55]. A sensor's false positive/negative rates also play a role. Smart home designers should help users navigate these nuances through sensible default policies and templates.

> Many sensors are susceptible to physical DoS attacks. Mitigations against physical DoS of sensors include redundant sensors of different types and carefully constructed default policies.

Audio- and vision-based sensing is vulnerable to many attacks. Basic audio-based sensing is susceptible to all types of attacks in Tables 2-3 [1], [17], [125], [163]. Visible-light camera sensing is also susceptible to all of these attacks, except for hardware attacks. For cameras, spoofing can be difficult, but replay attacks with photo or video input are feasible.

Existing defenses for sensing methods are insufficient for access control because they were designed for *authentication* instead. Most prior work on audio- and camera-based sensing lacks security analyses. The few that analyzed security focused on replay and spoofing attacks. Authentication assumes that unrecognized users are unauthorized. Thus, a large body of research has focused on preventing replay and spoofing attacks against audio- and camera-based sensing to avoid attackers from becoming recognized in this regard. A commonly proposed defense is to rely on secondary channels of information on the same device [136] or other devices [13], [165]. For example, 3D cameras (like Face ID on iPhones [8]) analyze depth information to deter simple, photo-based replay attacks. However, in access control, default-allow policies authorize *unrecognized* users, resulting in the possibility of physical DoS attacks. Therefore, for such

policies, an attacker can gain access by targeting one information channel (e.g., targeting an image's visual features by presenting a photo) and becoming unrecognizable to the system.

Existing defenses for audio- and camera-based sensing focus on attacks that compromise authentication, not access control. Attackers can exploit the default semantics of access-control policies to gain access, and physical DoS attacks become easier.

Physical adversarial examples can be effective for skilled, external attackers. For sensing methods that rely on machine learning, we noted whether they were susceptible to adversarial examples. Specifically, within the scope of context sensing and our threat model, we consider only physical adversarial examples. The attacker misleads the algorithms by adding physical perturbations to the environment or to themselves, instead of feeding data to the algorithms directly. Recent work has demonstrated the feasibility of such attacks for images [35], [36], [130] and audio [1], [17], [125]. Although some attacks require whitebox access to models, which is unrealistic for commodity smart home devices, blackbox attacks are also possible [35], [63], [86], [134], [142].

Internal attackers are less likely to use physical adversarial examples because they require substantial technical skills and resources to generate and test. Instead, they would use familiarity with the system to launch replay, spoofing, or physical DoS attacks to a similar end. However, if we consider *external opportunistic* attackers (e.g., a group of burglars) who do not have information about the victim, physical adversarial examples can be very effective. In fact, untargeted adversarial examples are strictly easier than targeted attacks. For example, attackers might want to attack face recognition on all security cameras in a neighborhood. In doing so, they can reuse and refine their adversarial examples.

Internal attackers may prefer replay, spoofing, and physical DoS attacks. Opportunistic external attackers may prefer adversarial examples.

7.2. Privacy

Except for cameras, cloud storage is not usually required when sensing contexts. We found that 79.8% ($n = 75$) of the examined sensing techniques do not require data storage on the cloud. Unfortunately, 10 of the 14 methods that use cameras do require cloud processing. Oftentimes, cloud storage is necessary for computationally intensive algorithms or large training datasets required to process video or image data online (e.g., neural networks for facial recognition). Privacy-preserving machine learning may alleviate this need. One approach is to protect the privacy of the training data. In federated learning [71], sensitive data stays local and only gradient updates are sent to the server. Another approach targets the inference stage by running the models locally or on the edge [49], [74]. Companies may prefer cloud storage because they can collect user data. Despite the risk of data exposure, some users may prefer cloud storage if it costs less.

Few sensing methods, often camera-based ones, require cloud processing. Federated learning or performing ML on the edge could obviate cloud processing.

Cameras/microphones are invasive but currently indispensable, thus necessitating privacy countermeasures. Users perceive age to be an important context for access control [55]. Unfortunately, most existing age-estimation methods rely on cameras or microphones, raising privacy concerns. Until privacy-preserving methods for age detection become possible, users may instead wish to record age while registering their identity during system setup.

Suppose cameras and microphones have to be used. To enhance bystanders' privacy, countermeasures against these sensing methods have been proposed, such as strategically blurring an image or jamming microphones with ultrasonic noise [22], [31], [158]. These proposals improve privacy, but also imperil the access control system, making it more likely to ignore attackers or confuse attackers with benign users. Therefore, detecting contexts with obfuscated sensor data may be another research direction. Raval et al. [117] proposed a utility-aware obfuscation mechanism for smartphone apps, which shows a promising road to privacy-preserving sensing in homes.

Privacy-invasive sensors may be essential. Privacy protections may weaken the access-control system.

Mismatch between required and collected data. Only 25 of 94 context-sensor pairs (26.6%) do not collect more data than needed to deduce the context. In contrast, 33.0% were *acceptable* and 40.4% were *poor* in our analysis. Most sensing methods marked as *poor* record unnecessary video or audio. Manufacturers typically rely on high-fidelity sensors, such as cameras or microphones, to sense contexts. This also happens when researchers use microphones on voice assistants or smartphones for ultrasonic-based sensing for their *wide availability*. While federated learning or edge computing may mitigate privacy concerns, they may also appear cryptic to the average user. These methods may therefore fail to alleviate user concerns about sensors inadvertently collecting invasive data. Future work should investigate effective means of communicating to users privacy considerations, such as using privacy labels [98] or visual indicators [33].

Competing interests between multiple stakeholders—manufacturers, researchers, designers, users—also contribute to this mismatch between the data required and the data collected. The designer might only want to know which room the user is occupying, but manufacturers and UbiComp researchers likely would want to collect information about the activity of the user in that room. Obtaining this extra knowledge enables the latter two parties to design and provide technology benefiting users in other aspects of their daily life. For the benefit of smart home owners and users, smart home systems and sensors should offer the ability to prioritize utility or privacy.

Most sensors collect more data than needed. User awareness and control of data collection is critical.

7.3. Access, Deployment, and Acceptability

Many sensing methods for authentication are not inclusive. Research in sensing and access control is

generally not inclusive to the elderly and groups with various disabilities. For example, the gait-sensing literature mostly does not consider people with walking disabilities. For inclusivity, contextual access-control systems must offer an array of sensors that allow *every* individual to authenticate an identity or person-specific context.

8. Discussion

Utilizing the framework: Smart home designers can utilize Tables 2–3 to identify the sensor(s) most suitable for identifying a given context based on a given home's access-control policies. They can also examine overlaps between contexts for each area of a home to identify opportunities for sensor reuse. Based on the purpose of sensing each context and where the sensors will be deployed, they can prioritize particular security, privacy, or usability criteria. For example, outdoor sensors for burglary protection should prioritize security over privacy. Sensors for a home entertainment system might want to prioritize usability, while those in a more private area of one's home should emphasize privacy.

Fusing sensors can increase security [13]. Our framework also helps designers identify a suite of sensors that vary in the attacks (replay, physical DoS) they resist.

Despite our effort to be as objective as possible in applying our framework to the 94 pairs of contexts and (prototype or commercial) sensors to create Tables 2–3, some of our decisions were necessarily subjective or open to debate. Furthermore, the presentation of a compact table cannot possibly capture the nuanced discussion underpinning why a single cell shows a particular decision. As such, we imagine our framework serving as the guiding principle for the expanded online version of Tables 2–3, with the online version serving as a living document that captures the nuances of evaluating particular sensors for particular contexts, welcoming contributions from the community. We have therefore seeded this expanded online version with summaries of our reasoning about why a specific rating is given in each cell. Anyone can review these notes and decide whether the reasoning is correct. If they are confident that the reasoning is wrong, or some important sensing method is missing, they are encouraged to create an issue or make a pull request on the associated GitHub repository.[4]

Implications for auditing: Auditing is an easy way to solve many security issues mentioned in Section 7.1. Mismatches between identities or situations detected by multiple sensors [13] could signal an attempted replay, spoofing, or physical DoS attack. Products like Samsung SmartThings support auditing by providing a recent history of sensor data readings.

However, auditing sensor data faces privacy obstacles. Access to sensor data may create privacy concerns between members of a household (e.g., parents and kids [145]). Also, from a security perspective, logs should not be deleted or changed. Mutable logs defeat the purpose of having a log, which makes the balance between security and privacy even harder to achieve. Moreover, the amount of data may be huge, which can cause usability issues as well. Smart home designers cannot expect users to spend days sifting through sensor logs, particularly since identifying problems might require manually analyzing correlations between sensors.

To ease these burdens and balance privacy and transparency, log data could be audited automatically to only highlight key results. For example, the system could notify users that it observed motion at home when the house appeared unoccupied, guiding the user through the perceived discrepancy. Future research should focus on transparency that minimally impacts privacy, which can potentially be applied to any intelligent system with a human in the loop.

Access control with partial information: Information collected by household sensors can be incomplete, leading to a faulty access-control system. The sensor could be compromised by malicious parties, disabled to protect bystander privacy, out of battery, or otherwise corrupted. Sensor fusion is critical for robust access control. Correlations between multiple sensors can help verify physical events even if any subset are compromised [13].

If sensor readings are merely missing, not corrupted, the home ought to aid the user in debugging the problem. If the home is alerted that data is missing intentionally, perhaps due to a sensor taken out of service to protect bystander privacy, it should let users know what access-control policies may be affected. The system could also provide suggestions for less intrusive replacement sensors. In case the system does not know why data is missing, the system would benefit from every access-control policy having a secure default behavior. In other words, missing data should not facilitate attacks.

The privacy implication of a general model. One thing we did not fully capture in our framework is the privacy implication of a general model. A general model is a machine learning model that is environment- and user-independent, which means there is no need for users to perform any training themselves. In our evaluation, we typically gave these types of sensing methods good privacy ratings because they should not require storage of personal data. However, as discussed in Section 5, we only considered the minimum data needed to run the sensing method for the privacy evaluation. The actual implementation may collect more than the minimum data required. In particular, manufacturers are incentivized to acquire more data to improve the accuracy of their models. Unfortunately, with general models, the adversary only needs the data to infer user activities, whereas in user-specific models they to know both the data and the model.

9. Conclusion

Contextual access control in homes is desirable, yet mostly unsupported. To bridge this gap, sensors can be used to detect contexts. However, they must defend against both expert and non-expert adversaries while respecting user privacy and usability. We proposed both a new adversarial model for context sensing in homes and a decision framework for evaluating potential sensors in terms of security, privacy, and usability. We applied this framework to common sensors through literature systematization, finding important trade-offs. We have made our framework and evaluations accessible in a public GitHub repository to facilitate updates and public discussion.

4. https://github.com/UChicagoSUPERgroup/eurosp21

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants CNS-1756011 and CCF-1837120. Earlence Fernandes was supported by the University of Wisconsin-Madison Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation.

References

[1] Hadi Abdullah, Washington Garcia, Christian Peeters, Patrick Traynor, Kevin R. B. Butler, and Joseph Wilson. Practical hidden voice attacks against speech and speaker recognition systems. In *Proc. NDSS*, 2019.

[2] Mike Addlesee, Rupert Curwen, Steve Hodges, Joe Newman, Pete Steggles, Andy Ward, and Andy Hopper. Implementing a sentient computing system. *Computer*, 34(8):50–56, 2001.

[3] Omar Alrawi, Chaz Lever, Manos Antonakakis, and Fabian Monrose. SoK: Security evaluation of home-based IoT deployments. In *Proc. IEEE S&P*, 2019.

[4] Amazon. Echo. https://www.amazon.com/echo.

[5] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J. Alex Halderman, Luca Invernizzi, Michalis Kallitsis, Deepak Kumar, Chaz Lever, Zane Ma, Joshua Mason, Damian Menscher, Chad Seaman, Nick Sullivan, Kurt Thomas, and Yi Zhou. Understanding the Mirai botnet. In *Proc. USENIX Security*, 2017.

[6] Apple. Healthkit. https://developer.apple.com/healthkit/.

[7] Apple. iBeacon. https://developer.apple.com/ibeacon/.

[8] Apple. Use Face ID on your iPhone or iPad Pro. https://support.apple.com/en-us/HT208109.

[9] Apple. About Touch ID advanced security technology, 2017. https://support.apple.com/en-us/HT204587.

[10] Noah Apthorpe, Pardis Emami-Naeini, Arunesh Mathur, Marshini Chetty, and Nick Feamster. You, me, and IoT: How internet-connected consumer devices affect interpersonal relationships. arXiv:2001.10608, 2020.

[11] Bharathan Balaji, Jian Xu, Anthony Nwokafor, Rajesh Gupta, and Yuvraj Agarwal. Sentinel: Occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings. In *Proc. SenSys*, 2013.

[12] Nata M. Barbosa, Joon S. Park, Yaxing Yao, and Yang Wang. "What if?" Predicting individual users' smart home privacy preferences and their changes. *PoPETS*, 2019(4):211–231, 2019.

[13] Simon Birnbach, Simon Eberz, and Ivan Martinovic. Peeves: Physical event verification in smart homes. In *Proc. CCS*, 2019.

[14] Alan Bränzel, Christian Holz, Daniel Hoffmann, Dominik Schmidt, Marius Knaust, Patrick Lühne, René Meusel, Stephan Richter, and Patrick Baudisch. GravitySpace: Tracking users and their poses in a smart room using a pressure-sensing floor. In *Proc. CHI EA*, 2013.

[15] A.J. Bernheim Brush, Jaeyeon Jung, Ratul Mahajan, and Frank Martinez. Digital neighborhood watch: Investigating the sharing of camera data amongst neighbors. In *Proc. CSCW*, 2013.

[16] A.J. Bernheim Brush, Bongshin Lee, Ratul Mahajan, Sharad Agarwal, Stefan Saroiu, and Colin Dixon. Home automation in the wild: Challenges and opportunities. In *Proc. CHI*, 2011.

[17] Nicholas Carlini and David A. Wagner. Audio adversarial examples: Targeted attacks on speech-to-text. In *Proc. DLS*, 2018.

[18] Z. Berkay Celik, Gang Tan, and Patrick D. McDaniel. IoTGuard: Dynamic enforcement of security and safety policy in commodity IoT. In *Proc. NDSS*, 2019.

[19] Jagmohan Chauhan, Yining Hu, Suranga Seneviratne, Archan Misra, Aruna Seneviratne, and Youngki Lee. BreathPrint: Breathing acoustics-based user authentication. In *Proc. MobiSys*, 2017.

[20] Shaxun Chen, Amit Pande, and Prasant Mohapatra. Sensor-assisted facial recognition: an enhanced biometric authentication system for smartphones. In *Proc. MobiSys*, 2014.

[21] Yuanying Chen, Wei Dong, Yi Gao, Xue Liu, and Tao Gu. Rapid: A multimodal and device-free approach using noise estimation for robust person identification. *IMWUT*, 1(3), 2017.

[22] Yuxin Chen, Huiying Li, Shan-Yuan Teng, Steven Nagels, Zhijing Li, Pedro Lopes, Ben Y. Zhao, and Haitao Zheng. Wearable microphone jamming. In *Proc. CHI*, 2020.

[23] Eun Kyoung Choe, Sunny Consolvo, Jaeyeon Jung, Beverly Harrison, Shwetak N. Patel, and Julie A. Kientz. Investigating receptiveness to sensing and inference in the home using sensor proxies. In *Proc. UbiComp*, 2012.

[24] Gabriele Civitarese, Stefano Belfiore, and Claudio Bettini. Let the objects tell what you are doing. In *Proc. UbiComp*, 2016.

[25] CO2Meter. https://www.co2meter.com/products/tim10-desktop-co2-temp-humidity-monitor.

[26] Cory Cornelius, Ronald Peterson, Joseph Skinner, Ryan Halter, and David Kotz. A wearable system that knows who wears it. In *Proc. MobiSys*, 2014.

[27] Heather Crawford. Adventures in authentication–position paper. In *Proc. SOUPS*, 2014.

[28] Ry Crist. Amazon and Google are listening to your voice recordings. Here's what we know about that. CNET, July 2019. https://www.cnet.com/how-to/amazon-and-google-are-listening-to-your-voice-recordings-heres-what-we-know/.

[29] Tamara Denning, Tadayoshi Kohno, and Henry M. Levy. Computer security and the modern home. *CACM*, 56(1):94–103, 2013.

[30] Hamdi Dibeklioglu, Fares Alnajar, Albert Ali Salah, and Theo Gevers. Combining facial dynamics with appearance for age estimation. *IEEE TIP*, 24(6), 2015.

[31] Mariella Dimiccoli, Juan Marín, and Edison Thomaz. Mitigating bystander privacy concerns in egocentric activity recognition with deep learning and intentional image degradation. *IMWUT*, 1(4), 2017.

[32] John R. Douceur. The Sybil attack. In *Proc. IPTPS*, 2002.

[33] Serge Egelman, Raghudeep Kannavara, and Richard Chow. Is this thing on?: Crowdsourcing privacy indicators for ubiquitous sensing platforms. In *Proc. CHI*, 2015.

[34] Pardis Emami-Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. Privacy expectations and preferences in an IoT world. In *Proc. SOUPS*, 2017.

[35] Kevin Eykholt, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Florian Tramèr, Atul Prakash, Tadayoshi Kohno, and Dawn Song. Physical adversarial examples for object detectors. In *Proc. WOOT*, 2018.

[36] Kevin Eykholt, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. Robust physical-world attacks on deep learning visual classification. In *Proc. CVPR*, 2018.

[37] Huan Feng, Kassem Fawaz, and Kang G. Shin. Continuous authentication for voice assistants. In *Proc. MobiCom*, 2017.

[38] Earlence Fernandes, Amir Rahmati, Kevin Eykholt, and Atul Prakash. Internet of Things security research: A rehash of old ideas or new intellectual challenges? *IEEE Security & Privacy*, 15(4):79–84, 2017.

[39] First Alert. Battery powered photo & ion smoke alarm. https://images-na.ssl-images-amazon.com/images/I/A1+UJjI+uPL.pdf.

[40] Fitbit Inc. How do I track my activity with my Fitbit device? https://help.fitbit.com/articles/en_US/Help_article/1785.

[41] Diana Freed, Jackeline Palmer, Diana Minchala, Karen Levy, Thomas Ristenpart, and Nicola Dell. "A stalker's paradise": How intimate partner abusers exploit technology. In *Proc. CHI*, 2018.

[42] Chuhan Gao, Kassem Fawaz, Sanjib Sur, and Suman Banerjee. Privacy protection for audio sensing against multi-microphone adversaries. *PoPETs*, 2019(2):146–165, 2019.

[43] Christine Geeng and Franziska Roesner. Who's in control? Interactions in multi-user smart homes. In *Proc. CHI*, 2019.

[44] Google. Google Nest Protect. https://store.google.com/us/product/nest_protect_2nd_gen.

[45] Google. Nest & Google - The best of Google. The best of Nest. https://store.google.com/category/google_nest.

[46] Google. Voice match and media on Google Home. https://support.google.com/googlenest/answer/7342711.

[47] Google Nest. Split-spectrum white paper. Technical report, June 2015.

[48] Google Nest Cam. Indoor - tech specs. https://store.google.com/us/product/nest_cam_specs.

[49] Sridhar Gopinath, Nikhil Ghanathe, Vivek Seshadri, and Rahul Sharma. Compiling KB-sized machine learning models to tiny IoT devices. In *Proc. PLDI*, 2019.

[50] GridEye. Infrared Array Sensor Grid-EYE: High Precision Infrared Array Sensor based on Advanced MEMS Technology. https://www.mouser.com/datasheet/2/315/ADI8000C65-1267019.pdf.

[51] Tobias Grosse-Puppendahl, Xavier Dellangnol, Christian Hatzfeld, Biying Fu, Mario Kupnik, Arjan Kuijper, Matthias R. Hastall, James Scott, and Marco Gruteser. Platypus - Indoor localization and identification through sensing electric potential changes in human bodies. In *Proc. MobiSys*, 2016.

[52] Tobias Grosse-Puppendahl, Sebastian Herber, Raphael Wimmer, Frank Englert, Sebastian Beck, Julian von Wilmsdorff, Reiner Wichert, and Arjan Kuijper. Capacitive near-field communication for ubiquitous interaction and perception. In *Proc. UbiComp*, 2014.

[53] Sidhant Gupta, Ke-Yu Chen, Matthew S. Reynolds, and Shwetak N. Patel. Lightwave: Using compact fluorescent lights as sensors. In *Proc. UbiComp*, 2011.

[54] Rakibul Hasan, David Crandall, Mario Fritz, and Apu Kapadia. Automatically detecting bystanders in photos to reduce privacy risks. In *Proc. IEEE S&P*, 2020.

[55] Weijia He, Maximilian Golla, Roshni Padhi, Jordan Ofek, Markus Dürmuth, Earlence Fernandes, and Blase Ur. Rethinking access control and authentication for the home Internet of Things (IoT). In *Proc. USENIX Security*, 2018.

[56] Alex Hern. Uber employees 'spied on ex-partners, politicians and Beyoncé', December 2016. https://www.theguardian.com/technology/2016/dec/13/uber-employees-spying-ex-partners-politicians-beyonce.

[57] Timothy W. Hnat, Erin Griffiths, Raymond Dawson, and Kamin Whitehouse. Doorjamb: Unobtrusive room-level tracking of people in homes using doorway sensors. In *Proc. SenSys*, 2012.

[58] Honeywell. 5853 Wireless Glassbreak Detector . https://www.security.honeywell.com/product-repository/5853.

[59] Honeywell. DT906 / DT907. https://www.security.honeywell.com/product-repository/dt906-dt907.

[60] Jason I. Hong and James A. Landay. An architecture for privacy-sensitive ubiquitous computing. In *Proc. MobiSys*, 2004.

[61] Chen-Yu Hsu, Rumen Hristov, Guang-He Lee, Mingmin Zhao, and Dina Katabi. Enabling identification and behavioral sensing in homes using radio reflections. In *Proc. CHI*, 2019.

[62] Yih-Chun Hu, Adrian Perrig, and David B. Johnson. Wormhole attacks in wireless networks. *IEEE J-SAC*, 24(2), 2006.

[63] Andrew Ilyas, Logan Engstrom, and Aleksander Madry. Prior convictions: Black-box adversarial attacks with bandits and priors. In *Proc. ICLR*, 2019.

[64] Timo Jakobi, Sameer Patil, Dave Randall, Gunnar Stevens, and Volker Wulf. It is about what they could do with the data: A user perspective on privacy in smart metering. *TOCHI*, 26(1), 2019.

[65] Timo Jakobi, Gunnar Stevens, Nico Castelli, Corinna Ogonowski, Florian Schaub, Nils Vindice, Dave W. Randall, Peter Tolmie, and Volker Wulf. Evolving needs in IoT control and accountability: A longitudinal study on smart home intelligibility. *IMWUT*, 2(4), 2018.

[66] Yunhan Jack Jia, Qi Alfred Chen, Shiqi Wang, Amir Rahmati, Earlence Fernandes, Z. Morley Mao, and Atul Prakash. ContexIoT: Towards providing contextual integrity to appified IoT platforms. In *Proc. NDSS*, 2017.

[67] Matthew Johnson and Frank Stajano. Usability of security management: Defining the permissions of guests. In *Proc. SPW*, 2006.

[68] Avinash Kalyanaraman, Dezhi Hong, Elahe Soltanaghaei, and Kamin Whitehouse. FormaTrack: Tracking people based on body shape. *IMWUT*, 1(3), 2017.

[69] Chris Karlof and David A. Wagner. Secure routing in wireless sensor networks: Attacks and countermeasures. *Ad Hoc Networks*, 1(2-3), 2003.

[70] Tiffany Hyun-Jin Kim, Lujo Bauer, James Newsome, Adrian Perrig, and Jesse Walker. Challenges in access right assignment for secure home networks. In *Proc. HotSec*, 2010.

[71] Jakub Konečnỳ, H. Brendan McMahan, Daniel Ramage, and Peter Richtárik. Federated optimization: Distributed machine learning for on-device intelligence. arXiv:1610.02527, 2016.

[72] Deepak Kumar, Kelly Shen, Benton Case, Deepali Garg, Galina Alperovich, Dmitry Kuznetsov, Rajarshi Gupta, and Zakir Durumeric. All things considered: An analysis of IoT devices on home networks. In *Proc. USENIX Security*, 2019.

[73] Denis Foo Kune, John D. Backes, Shane S. Clark, Daniel B. Kramer, Matthew R. Reynolds, Kevin Fu, Yongdae Kim, and Wenyuan Xu. Ghost talk: Mitigating EMI signal injection attacks against analog sensors. In *Proc. IEEE S&P*, 2013.

[74] Aditya Kusupati, Manish Singh, Kush Bhatia, Ashish Kumar, Prateek Jain, and Manik Varma. FastGRNN: A fast, accurate, stable and tiny kilobyte sized gated recurrent neural network. In *Proc. NeurIPS*, 2018.

[75] Nicholas D. Lane, Petko Georgiev, Cecilia Mascolo, and Ying Gao. Zoe: A cloud-less dialog-enabled continuous sensing wearable exploiting heterogeneous computation. In *Proc. MobiSys*, 2015.

[76] Gierad Laput, Karan Ahuja, Mayank Goel, and Chris Harrison. Ubicoustics: Plug-and-play acoustic activity recognition. In *Proc. UIST*, 2018.

[77] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *PACM HCI*, 2(CSCW), 2018.

[78] Patrick Lazik and Anthony Rowe. Indoor pseudo-ranging of mobile devices using ultrasonic chirps. In *Proc. SenSys*, 2012.

[79] Vassilios Lekakis, Yunus Basagalar, and Pete Keleher. Don't trust your roommate or access control and replication protocols in "home" environments. In *Proc. HotStorage*, 2012.

[80] Fan Li, Chunshui Zhao, Guanzhong Ding, Jian Gong, Chenxing Liu, and Feng Zhao. A reliable and accurate indoor localization method using phone inertial sensors. In *Proc. Ubicomp*, 2012.

[81] Tianxing Li, Qiang Liu, and Xia Zhou. Practical human sensing in the light. In *Proc. MobiSys*, 2016.

[82] Zhijing Li, Zhujun Xiao, Yanzi Zhu, Irene Pattarachanyakul, Ben Y. Zhao, and Haitao Zheng. Adversarial localization against wireless cameras. In *Proc. HotMobile*, 2018.

[83] Feng Lin, Chen Song, Yan Zhuang, Wenyao Xu, Changzhi Li, and Kui Ren. Cardiac scan: A non-contact and continuous heart-based user authentication system. In *Proc. MobiCom*, 2017.

[84] Hui Liu, Changyu Li, Xuancheng Jin, Juanru Li, Yuanyuan Zhang, and Dawu Gu. Smart solution, poor protection: An empirical study of security and privacy issues in developing and deploying smart home devices. In *Proc. IoT S&P*, 2017.

[85] Jian Liu, Cong Shi, Yingying Chen, Hongbo Liu, and Marco Gruteser. Cardiocam: Leveraging camera on mobile devices to verify users while their heart is pumping. In *Proc. MobiSys*, 2019.

[86] Yanpei Liu, Xinyun Chen, Chang Liu, and Dawn Song. Delving into transferable adversarial examples and black-box attacks. In *Proc. ICLR*, 2017.

[87] Edoardo Maggio. The facial recognition on Samsung's Galaxy Note 8 can be fooled with a photo. Business Insider, 2017. https://www.businessinsider.com/samsung-galaxy-note-8-facial-recognition-tricked-with-a-photo-2017-9.

[88] Nathan Malkin, Julia Bernd, Maritza Johnson, and Serge Egelman. "What can't data be used for?" Privacy expectations about smart TVs in the US. In *Proc. EuroUSEC*, 2018.

[89] Tsutomu Matsumoto, Hiroyuki Matsumoto, Koji Yamada, and Satoshi Hoshino. Impact of artificial "gummy" fingers on fingerprint systems. *OSCDT*, 2002.

[90] Tara Matthews, Kathleen O'Leary, Anna Turner, Manya Sleeper, Jill Palzkill Woelfer, Martin Shelton, Cori Manthorne, Elizabeth F. Churchill, and Sunny Consolvo. Stories from survivors: Privacy & security practices when coping with intimate partner abuse. In *Proc. CHI*, 2017.

[91] Vasilios Mavroudis, Shuang Hao, Yanick Fratantonio, Federico Maggi, Christopher Kruegel, and Giovanni Vigna. On the privacy and security of the ultrasound ecosystem. *PoPETs*, 2017(2):95–112, 2017.

[92] Michelle L. Mazurek, J. P. Arsenault, Joanna Bresee, Nitin Gupta, Iulia Ion, Christina Johns, Daniel Lee, Yuan Liang, Jenny Olsen, Brandon Salmon, Richard Shay, Kami Vaniea, Lujo Bauer, Lorrie Faith Cranor, Gregory R. Ganger, and Michael K. Reiter. Access control for home data sharing: Evaluating social acceptability. In *Proc. CHI*, 2010.

[93] Microsoft. Windows Hello: Discover facial recognition on Windows 10. https://www.microsoft.com/en-us/windows/windows-hello.

[94] Shaunak Mishra, Yasser Shoukry, Nikhil Karamchandani, Suhas N. Diggavi, and Paulo Tabuada. Secure state estimation against sensor attacks in the presence of noise. *IEEE TCNS*, 4(1):49–59, 2017.

[95] Reham Mohamed and Moustafa Youssef. Heartsense: Ubiquitous accurate multi-modal fusion-based heart rate estimation using smartphones. *IMWUT*, 1(3), September 2017.

[96] Adiyan Mujibiya and Jun Rekimoto. Mirage: Exploring interaction modalities using off-body static electric field sensing. In *Proc. UIST*, 2013.

[97] Dibya Mukhopadhyay, Maliheh Shirvanian, and Nitesh Saxena. All your voices are belong to us: Stealing voices to fool humans and machines. In *Proc. ESORICS*, 2015.

[98] Pardis Emami Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. Ask the experts: What should be on an IoT privacy and security label? In *Proc. IEEE S&P*, 2020.

[99] NAPCO. Napco adaptive dual microwave/PIR detector, 30x35 ft. (c-100ste). https://www.amazon.com/Napco-Adaptive-Microwave-Detector-C-100STE/dp/B0041X47EW.

[100] NAPCO. Napco's adaptive dual microwave/PIR detectors automatically adjust to their environment, minute by minute, for the ultimate false alarm immunity & reliability. https://napcosecurity.com/products/napco-detectors/.

[101] National Coordination Office for Space-Based Positioning, Navigation, and Timing. GPS: The Global Positioning System. https://www.gps.gov/.

[102] NETATMO. Smart indoor camera. https://www.netatmo.com/en-us/security/cam-indoor/specifications.

[103] Jared Newman. Internet-connected Hello Barbie doll can be hacked. PC World, December 2014. https://www.pcworld.com/article/3012220/internet-connected-hello-barbie-doll-can-be-hacked.html.

[104] Alfred Ng. Smart home tech can help evict renters, surveillance company tells landlords. CNET, October 2019. https://www.cnet.com/news/install-smart-home-tech-evict-renters-surveillance-company-tells-landlords/.

[105] Dat Tien Nguyen, So Ra Cho, Tuyen Danh Pham, and Kang Ryoung Park. Human age estimation method robust to camera sensor and/or face movement. *Sensors*, 15(9), 2015.

[106] Kazuya Ohara, Takuya Maekawa, Yasue Kishino, Yoshinari Shirai, and Futoshi Naya. Transferring positioning model for device-free passive indoor localization. In *Proc. UbiComp*, 2015.

[107] Robert J. Orr and Gregory D. Abowd. The smart floor: A mechanism for natural user identification and tracking. In *Proc. CHI EA*, 2000.

[108] Hongyu Pan, Hu Han, Shiguang Shan, and Xilin Chen. Mean-variance loss for deep age estimation from a face. In *Proc. CVPR*, 2018.

[109] Shijia Pan, Tong Yu, Mostafa Mirshekari, Jonathon Fagert, Amelie Bonde, Ole J. Mengshoel, Hae Young Noh, and Pei Zhang. FootprintID: Indoor pedestrian identification through ambient structural vibration sensing. *IMWUT*, 1(3), 2017.

[110] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The Cricket location-support system. In *Proc. MobiCom*, 2000.

[111] G. Qian, J. Zhang, and A. Kidané. People identification using floor pressure sensing and analysis. *IEEE Sensors Journal*, 10(9):1447–1460, 2010.

[112] Kun Qian, Chenshu Wu, Yi Zhang, Guidong Zhang, Zheng Yang, and Yunhao Liu. Widar2.0: Passive human tracking with a single Wi-Fi link. In *Proc. MobiSys*, 2018.

[113] Emilee Rader and Janine Slaker. The importance of visibility for folk theories of sensor data. In *Proc. SOUPS*, 2017.

[114] Tauhidur Rahman, Alexander T. Adams, Ruth Vinisha Ravichandran, Mi Zhang, Shwetak N. Patel, Julie A. Kientz, and Tanzeem Choudhury. Dopplesleep: A contactless unobtrusive sleep sensing system using short-range doppler radar. In *Proc. UbiComp*, 2015.

[115] Juhi Ranjan and Kamin Whitehouse. Object hallmarks: Identifying object users using wearable wrist sensors. In *Proc. UbiComp*, 2015.

[116] Aditya Singh Rathore, Weijin Zhu, Afee Daiyan, Chenhan Xu, Kun Wang, Feng Lin, Kui Ren, and Wenyao Xu. Sonicprint: A generally adoptable and secure fingerprint biometrics in smart devices. In *Proc. MobiSys*, 2020.

[117] Nisarg Raval, Ashwin Machanavajjhala, and Jerry Pan. Olympus: Sensor privacy through utility aware obfuscation. *PoPETS*, 2019(1):5–25, 2019.

[118] Steven Rosenbaum. Can you really 'ban' Google Glass? Forbes, June 2013. https://www.forbes.com/sites/stevenrosenbaum/2013/06/09/can-you-really-ban-google-glass/.

[119] Samsung. Galaxy S8 — S8+ - Security. https://www.samsung.com/global/galaxy/galaxy-s8/security/.

[120] Samsung. Motion sensor. https://www.lowes.com/pd/Samsung-Motion-Sensor/1000555661.

[121] Samsung. SmartThings, 2014. https://www.smartthings.com.

[122] Munehiko Sato, Rohan S. Puri, Alex Olwal, Yosuke Ushigome, Lukas Franciszkiewicz, Deepak Chandra, Ivan Poupyrev, and Ramesh Raskar. Zensei: Embedded, multi-electrode bioimpedance sensing for implicit, ubiquitous user recognition. In *Proc. CHI*, 2017.

[123] Florian Schaub, Rebecca Balebako, Adam L. Durity, and Lorrie Faith Cranor. A design space for effective privacy notices. In *Proc. SOUPS*, 2015.

[124] Stuart Schechter. The user IS the enemy, and (s)he keeps reaching for that bright shiny power button! In *Proc. HUPS*, 2013.

[125] Lea Schönherr, Katharina Kohls, Steffen Zeiler, Thorsten Holz, and Dorothea Kolossa. Adversarial attacks against automatic speech recognition systems via psychoacoustic hiding. In *Proc. NDSS*, 2019.

[126] Roei Schuster, Vitaly Shmatikov, and Eran Tromer. Situational access control in the Internet of Things. In *Proc. CCS*, 2018.

[127] James Scott, A.J. Bernheim Brush, John Krumm, Brian Meyers, Michael Hazas, Stephen Hodges, and Nicolas Villar. Preheat: Controlling home heating using occupancy prediction. In *Proc. UbiComp*, 2011.

[128] Mohammad Sedaaghi. A comparative study of gender and age classification in speech signals. *IJEEE*, 5, 03 2009.

[129] Souvik Sen, Dongho Kim, Stephane Laroche, Kyu-Han Kim, and Jeongkeun Lee. Bringing CUPID indoor positioning system to practice. In *Proc. WWW*, 2015.

[130] Mahmood Sharif, Sruti Bhagavatula, Lujo Bauer, and Michael K. Reiter. Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition. In *Proc. CCS*, 2016.

[131] Sierra Monitor Corporation. https://www.sierramonitor.com/flame-detector-3600-lb.

[132] Dan Simmons. BBC fools HSBC voice recognition security system. BBC, May 2017. https://www.bbc.com/news/technology-39965545.

[133] Milijana Surbatovich, Jassim Aljuraidan, Lujo Bauer, Anupam Das, and Limin Jia. Some recipes can do more than spoil your appetite: Analyzing the security and privacy risks of IFTTT recipes. In *Proc. WWW*, 2017.

[134] Fnu Suya, Jianfeng Chi, David Evans, and Yuan Tian. Hybrid batch attacks: Finding black-box adversarial examples with limited queries. In *Proc. USENIX Security*, 2020.

[135] Sheng Tan, Linghan Zhang, Zi Wang, and Jie Yang. Multitrack: Multi-user tracking and activity recognition using commodity wifi. In *Proc. CHI*, 2019.

[136] Di Tang, Zhe Zhou, Yinqian Zhang, and Kehuan Zhang. Face Flashing: A secure liveness detection protocol based on light reflections. In *Proc. NDSS*, 2018.

[137] Techamor. https://www.amazon.com/dp/B07BM1XWB8.

[138] Edison Thomaz, Vinay Bettadapura, Gabriel Reyes, Megha Sandesh, Grant Schindler, Thomas Plötz, Gregory D. Abowd, and Irfan Essa. Recognizing water-based activities in the home through infrastructure-mediated sensing. In *Proc. UbiComp*, 2012.

[139] Yuan Tian, Nan Zhang, Yue-Hsun Lin, XiaoFeng Wang, Blase Ur, Xianzheng Guo, and Patrick Tague. SmartAuth: User-centered authorization for the Internet of Things. In *Proc. USENIX Security*, 2017.

[140] B. Ugur Töreyin, R. Gokberk Cinbis, Yigithan Dedeoglu, and A. Enis Cetin. Fire detection in infrared video using wavelet analysis. *Opt. Eng.*, 46(6), 2007.

[141] Timothy Trippel, Ofir Weisse, Wenyuan Xu, Peter Honeyman, and Kevin Fu. WALNUT: Waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks. In *Proc. EuroS&P*, 2017.

[142] Chun-Chen Tu, Pai-Shun Ting, Pin-Yu Chen, Sijia Liu, Huan Zhang, Jinfeng Yi, Cho-Jui Hsieh, and Shin-Ming Cheng. Autozoom: Autoencoder-based zeroth order optimization method for attacking black-box neural networks. In *Proc. AAAI*, 2019.

[143] Sarah Underwood. Distinguishing identical twins. CACM, April 2018.

[144] Blase Ur, Jaeyeon Jung, and Stuart Schechter. The current state of access control for smart devices in homes. In *Proc. HUPS*, 2013.

[145] Blase Ur, Jaeyeon Jung, and Stuart Schechter. Intruders versus intrusiveness: Teens' and parents' perspectives on home-entryway surveillance. In *Proc. UbiComp*, 2014.

[146] Blase Ur, Elyse McManus, Melwyn Pak Yong Ho, and Michael L. Littman. Practical trigger-action programming in the smart home. In *Proc. CHI*, 2014.

[147] Edward J. Wang, Tien-Jui Lee, Alex Mariakakis, Mayank Goel, Sidhant Gupta, and Shwetak N. Patel. Magnifisense: Inferring device interaction using wrist-worn passive magneto-inductive sensors. In *Proc. UbiComp*, 2015.

[148] Ju Wang, Hongbo Jiang, Jie Xiong, Kyle Jamieson, Xiaojiang Chen, Dingyi Fang, and Binbin Xie. Lifs: Low human-effort, device-free localization with fine-grained subcarrier information. In *Proc. MobiCom*, 2016.

[149] Liang Wang, Tieniu Tan, Huazhong Ning, and Weiming Hu. Silhouette analysis-based gait recognition for human identification. *IEEE TPAMI*, 25(12), 2003.

[150] Qi Wang, Pubali Datta, Wei Yang, Si Liu, Adam Bates, and Carl A. Gunter. Charting the atack surface of trigger-action IoT platforms. In *Proc. CCS*, 2019.

[151] Qi Wang, Wajih Ul Hassan, Adam Bates, and Carl Gunter. Fear and logging in the Internet of Things. In *Proc. NDSS*, 2018.

[152] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures. In *Proc. MobiCom*, 2014.

[153] Roy Want, Andy Hopper, Veronica Falcão, and Jonathan Gibbons. The active badge location system. *ACM TIS*, 10(1):91–102, 1992.

[154] Venessa Wong. Burger King's new ad will hijack your Google Home. CNBC, 2017. https://www.cnbc.com/2017/04/12/burger-kings-new-ad-will-hijack-your-google-home.html.

[155] Chi-Jui Wu, Steven Houben, and Nicolai Marquardt. Eaglesense: Tracking people and devices in interactive spaces using real-time top-view depth-sensing. In *Proc. CHI*, 2017.

[156] Chen Yan, Hocheol Shin, Connor Bolton, Wenyuan Xu, Yongdae Kim, and Kevin Fu. SoK: A minimalist approach to formalizing analog sensor security. In *Proc. IEEE S&P*, 2020.

[157] Yaxing Yao, Justin Reed Basdeo, Smirity Kaushik, and Yang Wang. Defending my castle: A co-design study of privacy mechanisms for smart homes. In *Proc. CHI*, 2019.

[158] Hyunwoo Yu, Jaemin Lim, Kiyeon Kim, and Suk-Bok Lee. Pinto: Enabling video privacy for commodity IoT cameras. In *Proc. CCS*, 2018.

[159] Tianlong Yu, Vyas Sekar, Srinivasan Seshan, Yuvraj Agarwal, and Chenren Xu. Handling a trillion (unfixable) flaws on a billion devices: Rethinking network security for the Internet-of-Things. In *Proc. HotNets*, 2015.

[160] Jaeseok Yun. User identification using gait patterns on UbiFloorII. *Sensors*, 11(3), 2011.

[161] Eric Zeng, Shrirang Mare, and Franziska Roesner. End user security & privacy concerns with smart homes. In *Proc. SOUPS*, 2017.

[162] Eric Zeng and Franziska Roesner. Understanding and improving security and privacy in multi-user smart homes: A design exploration and in-home user study. In *Proc. USENIX Security*, 2019.

[163] Guoming Zhang, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. DolphinAttack: Inaudible voice commands. In *Proc. CCS*, 2017.

[164] Lefan Zhang, Weijia He, Jesse Martinez, Noah Brackenbury, Shan Lu, and Blase Ur. AutoTap: Synthesizing and repairing trigger-action programs using LTL properties. In *Proc. ICSE*, 2019.

[165] Linghan Zhang, Sheng Tan, and Jie Yang. Hearing your voice is not enough: An articulatory gesture based liveness detection for voice authentication. In *Proc. CCS*, 2017.

[166] Nan Zhang, Soteris Demetriou, Xianghang Mi, Wenrui Diao, Kan Yuan, Peiyuan Zong, Feng Qian, XiaoFeng Wang, Kai Chen, Yuan Tian, Carl A. Gunter, Kehuan Zhang, Patrick Tague, and Yue-Hsun Lin. Understanding IoT security through the data crystal ball: Where we are now and where we are going to be. arXiv:1703.09809, 2017.

[167] Weizhi Zhang, MI Sakib Chowdhury, and Mohsen Kavehrad. Asynchronous indoor positioning system based on visible light communications. *Opt. Eng.*, 53(4), 2014.

[168] Yang Zhang, Yasha Iravantchi, Haojian Jin, Swarun Kumar, and Chris Harrison. Sozu: Self-powered radio tags for building-scale activity sensing. In *Proc. UIST*, 2019.

[169] Yang Zhang, Chouchang Yang, Scott E. Hudson, Chris Harrison, and Alanson P. Sample. Wall++: Room-scale interactive and context-aware sensing. In *Proc. CHI*, 2018.

[170] Yu Zhang and Dit-Yan Yeung. Multi-task warped Gaussian process for personalized age estimation. In *Proc. CVPR*, 2010.

[171] Mingmin Zhao, Tianhong Li, Mohammad Abu Alsheikh, Yonglong Tian, Hang Zhao, Antonio Torralba, and Dina Katabi. Through-wall human pose estimation using radio signals. In *Proc. CVPR*, 2018.