To be published in the 2018 IET PETRAS Living in the Internet of Things: Cybersecurity of the IoT Conference

A capability-oriented approach to assessing privacy risk in smart home ecosystems

J. Sturgess, J. R. C. Nurse, and J. Zhao

Department of Computer Science, University of Oxford, Oxford, UK
{firstname.lastname}@cs.ox.ac.uk

Keywords: Internet-of-Things, Smart Home, Privacy, Risk

Abstract

Smart devices are increasingly ubiquitous; the multitude of risks they pose to user privacy continues to grow, but assessing such risks has proven difficult. In this paper, we discuss three factors which complicate the assessment of privacy risks in the context of the smart home. Firstly, smart devices are highly heterogeneous and hard to categorise, so top-down, taxonomy-oriented approaches to risk assessment do not fit well. Secondly, the threat landscape is vast, varied, and growing. Thirdly, the chief asset, personal information, is difficult to value—especially given that its value can be hugely affected by aggregation. To address these factors, we propose a novel, bottom-up approach in which the smart home ecosystem is reduced to its *data-collecting capabilities* (such as sensors and apps) and then privacy risk is assessed based on the information that the user *exposes*. We define a capability-oriented model which is system-neutral, extensible, and therefore well-suited to the fast-evolving nature of the smart home.

1 Introduction

A smart home is a home equipped with smart devices (also known as smart things or smart objects) [1], which are embedded systems providing residents with a means to monitor and control different features of the home. For instance, one smart device (a sensor) might trigger another (an actuator) to alter the state of a physical device (*e.g.*, a motion sensor triggering a door to be opened).

As smart home technology has evolved, smart devices have been networked to form smart home ecosystems. These ecosystems have enabled smart devices to combine efforts and provide benefits beyond just convenience [2], such as smart meters to improve energy efficiency and smart cameras triggered by sensors to improve home security. The smart devices within an ecosystem are typically networked to a central hub, which controls the information flow between them, configured and managed by the user via smartphone- or Web-based apps.

Smart devices are notoriously heterogeneous, largely unstandardised, and communicate over various common protocols (*e.g.*, Bluetooth LE, ZigBee, ZWave), leading to many concerns about security [3, 4, 5] and privacy [12]. Furthermore, the information that smart devices collect can yield insights about the user beyond what might reasonably be expected [6, 16]. However, as the popularity of smart home devices grows, it is becoming increasingly difficult for a user to identify and understand how much personal information is being exposed, and therefore the risks posed.

The number of distinct smart home ecosystems is increasing as companies seek to establish themselves in the market. Different ecosystems may treat compatible devices and their data differently, meaning the same device may pose different risks, or be vulnerable to different attack vectors, depending on its surrounding ecosystem.

Risk assessment in the smart home is complicated by the vast threat landscape, which scales with the number of devices and home users. Threat vectors arise from devices (which might join or leave dynamically), their apps, communications protocols, third party manufacturers and data controllers, and users. The smart home may include multiple users with varied levels of risk perception, any of whom may inadvertently leak information about the others. Personal information assets are intangible and difficult to protect; their value is hard to quantify and not always apparent (*e.g.*, it might be sold to insurers to profile users and tailor policies [7]). Moreover, many users are not keen on risk assessment, nor aware of the threats [15].

In this paper, we posit that in order to comprehensively assess privacy risk within a smart home ecosystem, we must do two things. Firstly, we must treat the ecosystem as an exhaustive list of the data-collecting capabilities of its constituent parts, rather than as a network of complex systems. Secondly, we must consider the risk to personal information in terms of exposure, rather than in terms of the likelihood and impact of particular attack vectors. By doing these two things, we strip away the complexities in both the ecosystem and the threat landscape, allowing us to focus on what personal information could be collected by the ecosystem and what risk it poses, rather than on the details of how or by whom.

The structure of this paper is as follows. In section 2, we review related work on smart device taxonomies, smart home threats

and assets (in the context of privacy), and risk models. In section 3, we survey market-leading smart home ecosystems to illustrate their differences in design and technology. In section 4, we define our capability-oriented model. In section 5, we draw on sections 2 and 3 to identify three factors which complicate smart home risk assessment, and then we discuss how our model works to address them.

2 Related Work

2.1 Smart Devices & Taxonomies

The definition and categorisation of smart devices continues to evolve as new technologies emerge. Alam *et al.* [9] offer a taxonomy of smart devices, broadly classifying each as either a sensor, a physiological device, or a multimedia device. However, this does not accommodate those smart devices which have multiple functionalities—*e.g.*, some smart bathroom scales[1] detect location, measure the user's weight, and display a local weather report, thereby placing the device in all three categories. In this taxonomy, knowing the category or categories into which a device is classified does not fully describe what that device might be capable of.

Lopez *et al.* [10] proposed the ISAND specification, characterising devices based on whether they have an Identity, Sensors, Actuators, or the ability to connect to a Network or make Decisions (*e.g.*, a smart car engine which can decide when its oil needs changing and sends a message to a display would need at least SAD-Smart functionality). This approach better reflects the plurality of functions that a smart device may have and, from a risk analysis perspective, allows us to associate a more accurate set of risks with a device based on which categories it is classified into. Recognition of identity is important, since information from a device with a unique identifier can be linked to a user and aggregated with other sources. However, this approach is still too coarse-grained to describe all smart devices accurately—*e.g.*, two devices, one with a microphone and the other with a motion sensor, would both be classified as having at least S-Smart functionality, but they pose different privacy risks, since the former might record a user's name or address, if spoken within its vicinity, whereas the latter could not.

2.2 Smart Home Threat Landscape

Ziegeldorf *et al.* [6] categorise privacy threats in the wider Internet-of-Things context and show that well-known threats such as identification, tracking, and profiling may be greatly exacerbated by smart technologies and that the following new threats emerge as a result:

- *privacy-violating interactions*, where personal information is transmitted over a public medium and leaked, such as in advertising a specialist product;
- *lifecycle transitions*, where personal information stored on a smart device may be accessible to another user, such as a new owner;
- *inventory attacks*, where the existence and details of the user's personal things are collected, potentially allowing for greater inference about that user; and
- *linkage*, where personal information from multiple sources is aggregated for increased insight, potentially with loss of context, no concern for correctness, and the risk of data de-anonymisation.

These new threats complicate risk assessment by introducing unexpected assets and subtle attack vectors. Risks to privacy in the smart home are varied [12, 13, 21], including many potential side channel and metadata attacks.

ENISA [14] identifies a lack of security and privacy considerations in the design stages as a concern in smart devices and provides a comprehensive threat landscape which includes spyware and adware in apps and devices, traffic monitoring, device leakage, and server compromise. In any smart home ecosystem, large amounts of exposed personal data will likely accrue in a small number of places, such as cloud servers or mobile devices; the risk posed by that data to user privacy is significant—whether stolen by an attacker, leaked accidentally, or sold by unscrupulous companies—because it is paired with the user and therefore ripe for aggregation. Roman *et al.* [21] advocate for privacy-by-design as a means to address these challenges.

Some smart devices are compatible with multiple smart home ecosystems, in some cases supporting different protocols when in different ecosystems [11] (*e.g.*, Philips Hue lights[2] are compatible with all of the ecosystems mentioned in §3). As such, the same smart device may face different threats depending on the ecosystem in which it is operating—*e.g.*, if the ecosystem relies on a remote server, a greater risk to data is posed by interception or server compromise; whereas, if the ecosystem stores data on a mobile device, a greater risk may be posed by physical theft or malware. Thus, risk assessment in the smart home must consider devices, their apps (many of which are provided by third parties and may be over-privileged [2]), and the ecosystem, as well as other sources of information.

2.3 Personal Information & Privacy Risk

Zang *et al.* [15] surveyed personal information leakage in mobile apps and identified sixteen types of personal information at risk (see Figure 1). The authors categorised each type as either personally identifiable information (PII), behavioural information, or location information to better communicate to users what is at risk.

Creese *et al.* [16] illustrate the varying degrees of ease with which one type of personal information may imply another within online social networks, and so how sharing information can yield greater insight than expected. They identify types of personal information relevant to social networking, including: ethnicity, facial biometrics, image location metadata, people tags, online groups, and places and times of social activities.

[1] Nokia WiFi smart scales, https://health.nokia.com/eu/en/scales (last accessed: December 2017).

[2] Philips Hue starter kit, https://www.philips.co.uk/c-p/8718696592946/hue-white-and-color-ambiance-starter-kit-e27 (last accessed: Dec. 2017).

Data category	Data type	Variation
PII	Address	street address, hometown
	Birthday	birthday (month, day), birth year
	Email	
	Gender	
	Name	first name, last name
	Password	
	Phone Info	
	Phone Number	
	ZIP code	
Behavior	Employment	job searches
	Friend	name, email, phone number
	Medical Info	diseases, medications, height, weight, diet, exercise
	Post	texts, chats, likes
	Search	clothing, groceries, locations
	Username	
Location	Location	current GPS location, city

Figure 1: Zang *et al.* [15] identified these sixteen types of personal information.

The wider implications of privacy risk in smart devices are considered by Arabo *et al.* [13], who identify many potential threats, including social engineering attacks (*e.g.*, phishing), social network attacks (*e.g.*, child-grooming and cyberbullying), and identity theft. These threats highlight the real-world importance of protecting personal information.

2.4 Smart Home Risk Modelling

Conventional wisdom in risk modelling has long been to consider the likelihood and impact of attacks on a given asset [17], with some more recent work applying this approach to risk assessment in the smart home [18, 19].

Denning *et al.* [20] present a scenario-driven, device-centric approach in which the risk posed to a given smart device is broken down into three components: the feasibility of the attack, the attractiveness of the device as a compromised platform, and the potential damage caused. Essentially, the first two express likelihood and the third expresses impact. The authors consider a wide range of assets, including emotional and societal well-being, but concede valuation is difficult.

Jacobsson *et al.* [23] propose a model with emphasis on smart home automation systems, which also relies on likelihood and impact estimation at its core. In this model, the user is identified as the greatest risk agent. The authors suggest that information classification may help to integrate privacy-specific issues into such models. In [19, 12], they explore the difficulty of estimating the value of personal information. The user is unlikely to know what information is available to be taken or what external information might be available to increase the value of what was taken (by aggregation). Furthermore, users value personal information differently—*e.g.*, Townsend *et al.* [22] investigated health monitoring in older users and found that most would trade their privacy for autonomy.

Djemame *et al.* [24] propose a risk assessment model for the wider cloud service context. However, they neglect the user perspective, which in the narrower context of the smart home is crucial, given the influence the user has.

3 Smart Home Topologies

Smart devices and hubs (or controllers) connect to one another or a remote, cloud server to share and receive data, enhance their effectiveness, and provide additional utility. These smart home ecosystems are networked differently depending on which hub is used; some examples follow.

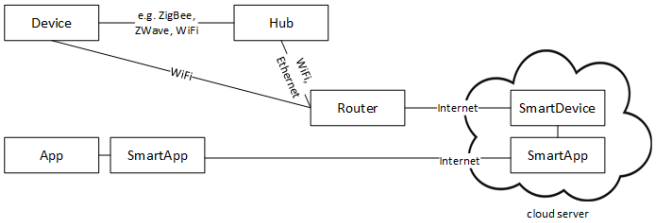

Figure 2: The Samsung SmartThings ecosystem.

In Samsung's SmartThings[3] ecosystem, each device uses a commonly supported protocol (*e.g.*, ZigBee, ZWave) to talk with a centralised hub, which remains in constant communication with a back-end cloud server over an SSL-protected protocol. On the server, each supported device has a corresponding *SmartDevice* (a software wrapper encapsulating it) and a *SmartApp* (which controls the SmartDevice). To manage a device, the user installs the SmartApp on his mobile device, which communicates with the corresponding SmartApp on the server. This, in turn, interacts with the SmartDevice, which communicates with the physical device via the hub (see Figure 2). Fernandez *et al.* [2] identified a number of security concerns in the SmartThings framework, such as rife over-privilege in SmartApps, many of which have since been addressed.

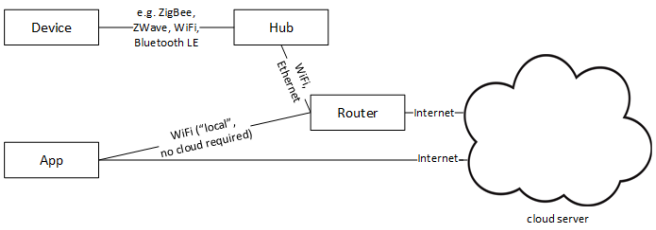

Figure 3: The Wink Hub ecosystem.

Wink Hub's[4] ecosystem is similar to SmartThings; each device communicates over a common protocol with the hub, thence to a cloud back-end over an SSL-protected protocol. To manage a device, the user installs its app on his mobile device and communicates either via the cloud server or, for devices supporting

[3] SmartThings Developer Documentation, http://docs.smartthings.com/en/latest/architecture (last accessed: December 2017).

[4] Wink Hub FAQs and Product Support, https://www.wink.com/help/faq (last accessed: December 2017).

the 'local' feature, via the router by connecting directly over the home network (see Figure 3). Another product, Wink Relay, offers a dedicated control panel to manage devices. With regard to risk, Veracode [25] found that communications between the Wink Hub and server were vulnerable to MITM attack due to a lack of TLS certificate validation; the authors also noted that a full breach of the Samsung or Wink services would allow an attacker access to the state of all products and services paired with every user.

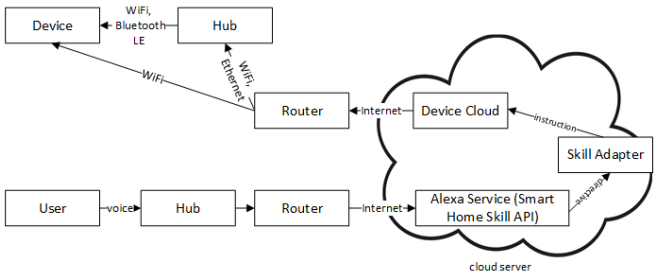

Figure 4: The Amazon Alexa ecosystem.

Amazon's Alexa[5] ecosystem implements a voice-controlled hub which is constantly connected to a cloud back-end. Whenever a voice command (a statement preceded by "Alexa") is detected, the recording is passed to the Alexa service on the cloud server where it is parsed by a programmable *skill* [26] to identify the relevant device and the action being requested (and user authentication). The derived *directive* is then passed to the user's device cloud to execute the action (see Figure 4). Skills can be programmed by third parties; this design allows cross-compatibility with other ecosystems—*e.g.*, SmartThings hubs and devices can be controlled by voice commands using Alexa, provided the necessary skills are installed. A similar ecosystem, Google Home[6], collects data ostensibly to improve its services [27]; Amazon Alexa currently does not, although this may change in the near future [28]. If transcripts are to be shared, this poses a risk to user privacy by exposing what devices are used and when; furthermore, the system detects voice commands by identifying trigger words, which may be spoken or detected erroneously, leading to information leakage.

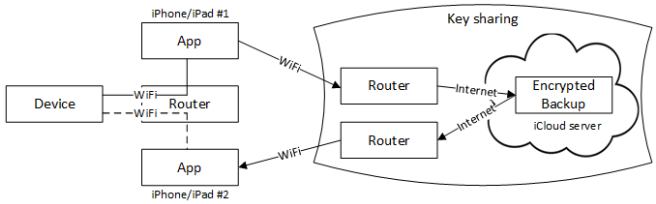

Figure 5: The Apple HomeKit ecosystem.

Apple's HomeKit[7] ecosystem has a decentralised topology and does not require a hub or back-end server. Instead, each device may be controlled directly by its app, installed on the user's mobile device, over a Station-to-Station protocol. Multiple mobile devices may exchange keys to share control of devices by transferring an encrypted backup via a cloud server (see Figure 5). Apple's upcoming HomePod[8] hub adds optional centralisation to the ecosystem, although technical details are not yet available at time of writing.

4 A Model for Assessing Privacy Risk in Smart Home Ecosystems

4.1 Model Definition

We define a *component* of a smart home ecosystem to be any smart device or hub connected to that ecosystem, bounded by and excluding the home router. We define a *capability* to be any functionality provided by a component or its accompanying app (*e.g.*, the ability to measure temperature, record audio, or take medical readings). We can therefore reduce any smart home ecosystem to a list of its components, and thence to a list of its capabilities.

We define a *risk profile* for a given capability to be a comprehensive risk assessment of that capability, identifying the risk it poses to each type of personal information with which we are concerned.

To assess privacy risk in the smart home, we assume that any personal information exposed to the ecosystem is available to an adversary. This enables us to strip away the complexities of the threat landscape and focus on the types of personal information the ecosystem is capable of collecting. We do so by first reducing it to a list of its data-collecting capabilities; since each capability has a corresponding risk profile, we can combine these risk profiles to get an overall privacy risk assessment for that ecosystem. The model is shown in Figure 6.

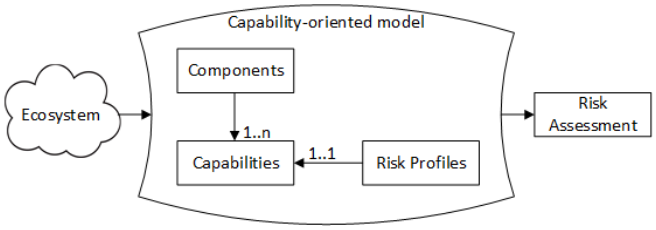

Figure 6: A capability-oriented model for assessing privacy risk in smart home ecosystems.

4.2 Types of Capability

To establish an exemplar list of capabilities, we surveyed commercial components and compiled a bottom-up list of common data-collecting capabilities.

[5] Alexa, https://developer.amazon.com/alexa (last accessed: Dec. 2017).
[6] Google Home, https://store.google.com/product/google_home (last accessed: December 2017).
[7] iOS Security, https://www.apple.com/business/docs/iOS_Security_Guide.pdf (last accessed: December 2017).
[8] Apple HomePod, https://www.apple.com/uk/homepod (last accessed: December 2017).

For instance, Samsung's SmartThings[9] and Amazon's Echo[10] are two popular hubs. The former contains a hub, a motion sensor, a presence sensor, a power usage sensor, and a multi-sensor (to monitor various environmental parameters); a companion app[11] is required to control the hub and connected devices, which requires permission to access one's identity, contacts, location, files, network, and camera. The latter has a built-in microphone and an app[12] requiring similar permissions, as well as access to one's phone and microphone. We treat each sensor and permission as a capability, since each can collect personal information. Both systems require the user to create an account with an email address, password, and other optional fields. We treat this requirement as another capability, since users will generally submit genuine information.

To sample the range of smart home technology, we examined three popular types of consumer device. Smart televisions[13] typically have an integrated Web browser (to support online streaming services), a microphone, and require an account and manual activation. Smart refrigerators[14] have environment and energy usage sensors and sometimes an app[15]. Smart scales[16] sense medical or health information, are manually activated (insofar as the user needs to be present or connected to it), and sometimes have a GPS-backed weather display.

By examining these and other components, we compiled an example list of capabilities, as shown in Table 1.

Capability	Notes
Microphone	records audio
Camera	records video and audio
Presence Sensor	senses presence, motion, *etc*.
Environment Sensor	senses temperature, light, *etc*.
Consumption Sensor	senses usage of water, power, *etc*.
GPS	records location
Medical/Health Reader	senses medical/health information
Timed Activation	turned on/off by a timer or remote user
Manual Activation	turned on/off by a present user
Weather Display	backed by GPS
Integrated Web Browser	records browsing history
Integrated Television	records viewing history
User Account	email and password required
App: Access to Identity	permission required
App: Access to Contacts	permission required
App: Access to Microphone	permission required
App: Access to Camera	permission required
App: Access to Location	permission required
App: Access to Network	permission required
App: Access to Files	permission required
App: Access to Phone	permission required

Table 1: A list of common data-collecting capabilities possessed by smart home components.

4.3 Types of Personal Information

By considering the capabilities identified in Table 1 and the types of personal information they may collect, and noting our findings from §2.3, we compiled an example list of types of personal information, as shown in Table 2.

Personal Information	Notes
Address	location of home
Birthday/Age	
Email	
Gender	
Phone Number	
Name/Alias	name or associable usernames
Password	
Friends/Associates	
Ethnicity	
Employment Details	employment status or job title
Religious/Political Beliefs	
Interests/Hobbies	
Routine	
Medical/Health Details	
Spending Habits	
Banking/Card Details	
Search History	
Social Media Activity	
User Presence	whether user is at home and awake
User Location	location of user if not at home

Table 2: A list of types of personal information which may be collected by a smart home ecosystem.

4.4 Risk Profiles

For each capability in Table 1, we created a risk profile expressing the severity of risk posed by that capability to each type of personal information in Table 2.

We used 'H', 'M', and 'L' to indicate a high, medium, or low severity of risk, respectively, and '-' to indicate no or negligible risk. The following rules were observed to classify severities:

- *low*: an indirect risk with a single threat vector (*e.g.*, a microphone poses a low risk to medical information, since it would only be collected if a conversation on this topic were recorded);
- *medium*: an indirect risk with multiple threat vectors (*e.g.*, a microphone poses a medium risk to gender, age, and routine information, since they would be collected or strongly inferred if any conversation on a range of topics were recorded);
- *high*: a direct risk, where the capability is either designed to collect this information or requires its collection for its intended functionality (*e.g.*, a microphone poses a high risk to detecting if a user is present).

We present our risk profiles in matrix form in Figure 7.

[9] SmartThings Starter Kit, http://www.samsung.com/uk/smartthings/kit-f-str-kit-uk (last accessed: December 2017).

[10] Amazon Echo, https://www.amazon.co.uk/Amazon-SK705DI-Echo-Black/dp/B01GAGVIE4 (last accessed: December 2017).

[11] SmartThings app, https://play.google.com/store/apps/details?id=com.smartthings.android (last accessed: December 2017).

[12] Amazon Alexa app, https://play.google.com/store/apps/details?id=com.amazon.dee.app (last accessed: December 2017).

[13] LG TV, http://www.lg.com/uk/tvs/lg-OLED65W7V (l. a.: Dec. 2017); Sony TV, https://www.sony.co.uk/electronics/4k-tvs (l. a.: Dec. 2017).

[14] Samsung fridge, http://www.samsung.com/us/explore/family-hub-refrigerator/refrigerator (last accessed: December 2017); Siemens fridge, http://www.siemens-home.bsh-group.com/uk/productlist/cooling/fridges/freestanding-fridges/KS36VAI41G (last accessed: December 2017).

[15] Samsung Family Hub, https://play.google.com/store/apps/details?id=com.samsung.familyhub (last accessed: December 2017).

[16] Salter scales, http://www.salterhousewares.co.uk/salter-max-electronic-digital-bathroom-scales-silver.html (last accessed: December 2017); Nokia scales, https://health.nokia.com/eu/en/scales (l. a.: Dec. 2017).

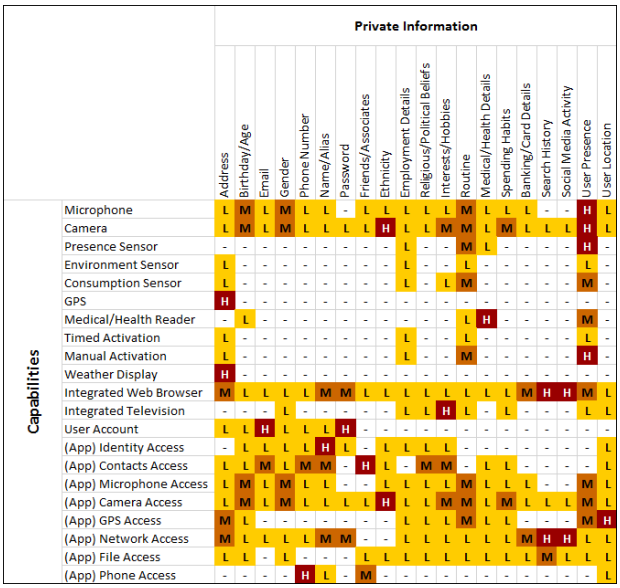

Figure 7: A privacy risk matrix. Each row is a risk profile for a capability, expressing the risk it poses to each type of personal information. We use red, orange, and yellow to represent high, medium, and low severities, respectively.

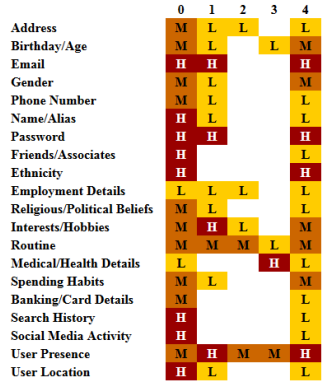

Figure 8: A risk assessment visualisation showing only the most severe risks per component (indexed along the top).

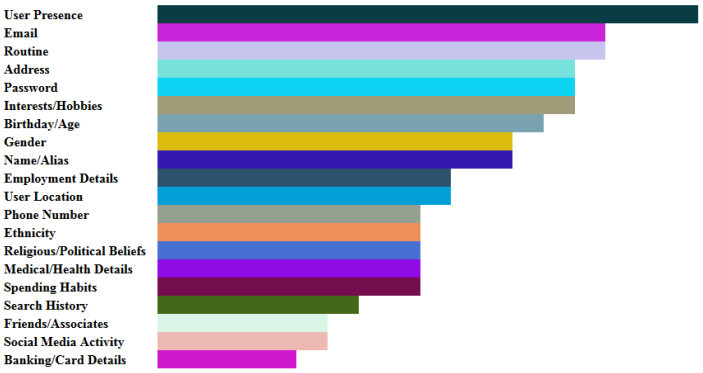

Figure 9: A risk assessment visualisation showing cumulative risk per type of personal information. We use arbitrary colours here only to distinguish between types.

4.5 Risk Assessment Visualisations

The flexibility of our model allows for risk profiles to be combined in different ways for different purposes. For instance, Figures 8 and 9 demonstrate two risk assessment visualisations for the same smart home ecosystem.

In Figure 8, each row represents a type of personal information and each column a separate component in the ecosystem. Each result shows only the most severe risk posed by that component, based on its capabilities, to that type of personal information. This visualisation enables the user to compare the risks posed by each component. It could be used to identify risky components or to see the impact of adding new components.

In Figure 9, we have mapped the values {3, 2, 1, 0} to the severities {H, M, L, -} and then, for each type of personal information, displayed the sum of the risks posed to that type by the ecosystem. This visualisation conveys which types of personal information are at greatest risk in the ecosystem and should only be used in the context of raising user awareness. (It is not intended to imply that, for instance, two low-severity risks strictly equate to one medium-severity risk, nor that the largest bar poses the most significant impact.)

5 Discussion

From our findings in sections 2 and 3, we identify three factors which complicate privacy risk assessment in the smart home. Firstly, smart devices are highly heterogeneous and therefore difficult to categorise; attempts to classify smart devices at a component level are too coarse-grained and fail to fully encapsulate all of the personal information that the ecosystem can collect. Secondly, smart home ecosystems are diverse and have a vast and complicated threat landscape, with many (often unstandardised) technologies and multiple users capable of leaking personal information about each other. Thirdly, personal information assets are intangible and difficult to value. Given that personal information is stored in only a small number of places in smart homes, densely and inter-twined, it is hard to ascertain how much of it is obtained by the adversary in a given attack. This means it is difficult to gauge the value of the information obtained, since (i) *internal* aggregation between separate pieces of obtained information potentially increases each piece's value, and (ii) *external* aggregation potentially yields greater results owing to the increased number of possible aggregation search keys obtained.

To address these factors, in our model we make the assumption that any personal information which is able to be collected by the ecosystem is available to the adversary and therefore at risk. While viewing a smart home ecosystem as a list of its capabilities does not fully describe the ecosystem, it does fully encapsulate what information the ecosystem is able to collect. Thus, our capability-oriented model enables one to consider all of the personal information that an ecosystem might collect about the user and then comprehensively assess the privacy risk

it poses. By approaching privacy risk assessment in this way, we address the first and second of the aforementioned factors, because we avoid both the need to classify smart devices (we only need to identify their data-collecting capabilities) and the need to contend with the intricacies of the threat landscape.

We posit that conventional risk assessment models—those which consider the likelihood and impact of attacks on a given asset—become unsuitable in the context of privacy risk in smart home ecosystems. Given the intangible nature of personal information and the difficulty one might have in attributing how an adversary acquired it, it should not matter whether it is stolen, leaked, or sold—only whether it is collected in the first place.

By stripping away complexity, our model identifies the personal information exposed to the ecosystem and enables the user to consider the value of that information both separately and as a whole. The user may assume that all pieces of information can be aggregated together to yield greater value. Thus, we address the third factor by presenting a more complete privacy risk picture of all the information which could be obtained by the adversary, making the variability posed by internal aggregation more estimable and reducing the uncertainty posed by external aggregation. We do not consider external aggregation any further, since determining precisely how much personal information is known would not be practically possible.

Our adversary is intentionally vague. We consider as an adversary any actor that collects personal information about the user, legally or illegally, and for any purpose, immediate or future; this therefore includes both malicious attackers and arguably benign actors (*e.g.*, service providers and device manufacturers), as even legitimate uses of personal information can leave users feeling that their privacy has been violated (*e.g.*, [7]).

6 Conclusion & Future Work

We have presented a model for assessing privacy risk in a smart home ecosystem. The novel part of this model is that it reduces an ecosystem to its data-collecting capabilities, then separately assesses the privacy risk posed by each using pre-defined risk profiles. This ensures that no data-collecting mechanism is overlooked and therefore fully encapsulates what personal information the ecosystem is able to collect. We identified three factors which complicate privacy risk assessment in the smart home and showed that our model addresses them.

A limitation of this work is that it does not consider personal information known to external sources, because assessing the amount and compounding value of this information would be impractical.

Future work might include the development of a tool to implement our model. This would enable the user to model his smart home and compare components based on the risks they pose, with the intention that it have a positive impact on user behaviour (Harkous *et al.* [8] showed similar results). It could also highlight common threats, such as identity theft, based on contributory factors identified in the ecosystem (risks to address, birthday, name, *etc.*), which might have a greater impact on users given its real-world relevance. Moreover, we envision that our risk profiles could be improved by incorporating crowd-sourced information to better elicit the risks posed in each case.

References

[1] V. Ricquebourg, D. Menga, D. Durand, B. Marhic, L. Delahoche, and C. Loge. "The Smart Home Concept: our immediate future", *IEEE International Conference on E-Learning in Industrial Electronics* (2006).

[2] E. Fernandes, J. Jung, and A. Prakash. "Security Analysis of Emerging Smart Home Applications", *IEEE Symposium on Security and Privacy* (2016).

[3] T. Simpson. "Securing a Heterogeneous Internet-of-Things", https://now.avg.com/securing-a-heterogeneous-internet-of-things (last accessed: December 2017).

[4] B. Violino. "IoT pushes IT security to the brink", http://www.csoonline.com/article/3081228/internet-of-things/iot-pushes-it-security-to-the-brink.html (last accessed: December 2017).

[5] J. Qi, A. V. Vasilakos, J. Wan, J. Lu, and D. Qiu. "Security of the Internet of Things: perspectives and challenges", *Wireless Networks*, **Volume 20, Issue 8**, pp. 2481-2501 (2014).

[6] J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle. "Privacy in the internet of things: Threats and challenges", *Security and Communication Networks* (2014).

[7] Allianz eBroker. "Home is where the smart is", https://www.allianzebroker.co.uk/news-and-insight/news/home-is-where-the-smart-is.html (last accessed: December 2017).

[8] H. Harkous, R. Rahman, B. Karlas, and K. Aberer. "Data-driven Privacy Indicators", *Workshop on Privacy Indicators, Twelfth Symposium on Usable Privacy and Security* (2016).

[9] M. R. Alam, M. B. I. Reaz, and M. A. M. Ali. "A Review of Smart Homes—Past, Present, and Future", *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, **Volume 42, Issue 6** (2012).

[10] T. S. Lopez, D. C. Ranasinghe, B. Patkai, and D. McFarlane. "Taxonomy, technology and applications of smart objects", *Information Systems Frontiers*, **Volume 13, Issue 2**, pp. 281-300 (2011).

[11] OVO Energy. "How do my smart meters communicate?", https://www.ovoenergy.com/ovo-answers/topics/smart-technology/smart-meters/how-do-my-smart-meters-communicate.html (last accessed: December 2017).

[12] J. Bugeja, A. Jacobsson, and P. Davidsson. "On Privacy and Security Challenges in Smart Connected Homes", *European Intelligence and Security Informatics Conference (EISIC)* (2016).

[13] A. Arabo, I. Brown, and F. El-Moussa. "Privacy in the Age of Mobility and Smart Devices in Smart Homes", *International Conference on Privacy, Security, Risk, and Trust (PASSAT)* (2012).

[14] D. Barnard-Wills, L. Marinos, and S. Portesi. "Threat Landscape and Good Practice Guide for Smart Home and Converged Media", *ENISA* (2014).

[15] J. Zang, K. Dummit, J. Graves, P. Lisker, and L. Sweeney. "Who Knows What About Me? A Survey of Behind the

Scenes Personal Data Sharing to Third Parties by Mobile Apps", https://techscience.org/a/2015103001 (last accessed: December 2017).

[16] S. Creese, M. Goldsmith, J. R. C. Nurse, and E. Phillips. "A Data-Reachability Model for Elucidating Privacy and Security Risks Related to the Use of Online Social Networks", *IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications* (2012).

[17] NIST. "Guide for Applying the Risk Management Framework to Federal Information Systems", *Special Publication 800-37* (2010).

[18] J. R. C. Nurse, A. Atamli, and A. Martin. "Towards a Usable Framework for Modelling Security and Privacy Risks in the Smart Home", *Human Aspects of Information Security, Privacy, and Trust*, pp. 255-267 (2016).

[19] A. Jacobsson and P. Davidsson. "Towards a Model of Privacy and Security for Smart Homes", *IEEE 2nd World Forum on Internet of Things (WF-IoT)* (2015).

[20] T. Denning, T. Kohno, and H. M. Levy. "Computer Security and the Modern Home", *Communications of the ACM*, **Volume 56, Issue 1**, pp. 94-103 (2013).

[21] R. Roman, P. Najera, and J. Lopez. "Securing the Internet of Things", *IEEE Computer*, **Volume 44, Issue 9**, pp. 51-58 (2011).

[22] D. Townsend, F. Knoefel, and R. Goubran. "Privacy Versus Autonomy: A Tradeoff Model for Smart Home Monitoring Technologies", *Annual International Conference of the IEEE, Engineering in Medicine and Biology Society* (2011).

[23] A. Jacobsson, M. Boldt, and B. Carlsson. "A risk analysis of a smart home automation system", *Future Generation Computer Systems*, **Volume 56**, pp. 719-733 (2016).

[24] K. Djemame, D. J. Armstrong, M. Krian, and M. Jiang. "A risk assessment framework and software toolkit for cloud service ecosystems", *Proceedings of the 2nd Int. Conf. on Cloud Computing, GRIDs, and Visualization* (2011).

[25] Veracode. "The Internet of Things: Security Research Study" (2015).

[26] Amazon. "Understanding the Smart Home Skill API", https://developer.amazon.com/public/solutions/alexa/alexa-skills-kit/overviews/understanding-the-smart-home-skill-api (last accessed: December 2017).

[27] Google. "Data security & privacy on Google Home", https://support.google.com/googlehome/answer/7072285 (last accessed: December 2017).

[28] R. LeFebvre. "Amazon may give developers your private Alexa transcripts", https://www.engadget.com/2017/07/12/amazon-developers-private-alexa-transcripts (last accessed: December 2017).