The Privacy Implications of Cyber Security Systems: A Technological Survey

Article *in* ACM Computing Surveys · February 2018

DOI: 10.1145/3172869

The Privacy Implications of Cyber Security Systems: A Technological Survey

ERAN TOCH, Tel Aviv University
CLAUDIO BETTINI, University of Milan
EREZ SHMUELI and LAURA RADAELLI, Tel Aviv University
ANDREA LANZI, University of Milan
DANIELE RIBONI, University of Cagliari
BRUNO LEPRI, Fondazione Bruno Kessler

Cyber-security systems, which protect networks and computers against cyber attacks, are becoming common due to increasing threats and government regulation. At the same time, the enormous amount of data gathered by cyber-security systems poses a serious threat to the privacy of the people protected by those systems. To ground this threat, we survey common and novel cyber-security technologies and analyze them according to the potential for privacy invasion. We suggest a taxonomy for privacy risks assessment of information security technologies, based on the level of data exposure, the level of identification of individual users, the data sensitivity and the user control over the monitoring, and collection and analysis of the data. We discuss our results in light of the recent technological trends and suggest several new directions for making these mechanisms more privacy-aware.

CCS Concepts: • **Security and privacy** → **Privacy protections**; *Malware and its mitigation*; *Intrusion detection systems*; *Information flow control*; *Firewalls*; • **Networks** → **Network privacy and anonymity**;

Additional Key Words and Phrases: Information security, privacy, system monitoring, network surveillance, privacy-preserving methods

ACM Reference format:
Eran Toch, Claudio Bettini, Erez Shmueli, Laura Radaelli, Andrea Lanzi, Daniele Riboni, and Bruno Lepri. 2018. The Privacy Implications of Cyber Security Systems: A Technological Survey. *ACM Comput. Surv.* 51, 2, Article 36 (February 2018), 27 pages.
https://doi.org/10.1145/3172869

This work is part of the Israeli-Italian Scientific and Technological Cooperation program, supported by the Ministry of Science, Technology and Space, Israel under grant 3-12288, and by the Ministry of Foreign Affairs and International Cooperation, Italy.

Author's addresses: E. Toch, E. Shmueli, and L. Radaelli are from the Department of Industrial Engineering, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel; emails: {erant, shmueli, radaelli}@tau.ac.il; C. Bettini and A. Lanzi are from the Dipartimento di Informatica, Università degli Studi di Milano, Via Comelico, 39, I-20135 Milano, Italy; emails: {claudio.bettini, andrea.lanzi}@unimi.it; D. Riboni is from the Dipartimento di Matematica e Informatica, Università degli Studi di Cagliari, Via Ospedale, 72, I-09124 Cagliari, Italy; email: riboni@unica.it; B. Lepri is from Fondazione Bruno Kessler, Via Sommarive, 18 - POVO, 38123 Trento, Italy; email: lepri@fbk.eu.

https://doi.org/10.1145/3172869

1 INTRODUCTION

In recent years, governments and corporations have increasingly relied on cyber-security systems to protect against increasing threats on networks, devices, and organizational and personal information. These systems prevent adversaries from breaking into networks and devices, from sabotaging digital activity, and from accessing private information. At the same time, by monitoring networks and computing devices, cyber-security systems ultimately affect individuals' privacy. Systems in domains such as intrusion detection, malware detection, data leakage prevention, and phishing identification regularly monitor network traffic, device use, and personal communications. In many cases, the monitoring system can trace the identities of users and access sensitive information. For instance, many enterprise cyber-security systems monitor IP addresses that can be easily traced back to a particular individual. Moreover, the user's device identification on mobile devices is often accessed by cyber-security applications. Therefore, while cyber-security mechanisms protect individuals from attacks from hackers and other third-party adversaries, they also create new vulnerabilities for privacy violation from the entity that runs the cyber-security system. These vulnerabilities can be realized if the security systems themselves are compromised,[1] if insiders make use of this information, or if the personal data are used contrary to the expectations of end users.[2]

The increasing threat of computer attacks and the intrusiveness of cyber-security mechanisms present policymakers and technology developers with the difficult challenge of balancing security risks against privacy and civil liberties concerns (Tene 2014; Landau 2014). The fact that many national cyber-security policies require the sharing of the detailed information of attack logs and other types of information necessitates an urgent understanding of the privacy risks related to cyber-security (Sales 2013; Nolan 2015). Privacy concerns are among the reasons why employees switch to their personal devices (e.g., smartphones and portable computers) to perform work-related activities (Pfleeger et al. 2014) and home-users turn away from some anti-virus applications (Warkentin and Willison 2009). Therefore, understanding and solving privacy threats is crucial, as those threats can reduce the acceptance and usage of cyber-security systems by organizations and individuals, leading to increased number of threats for everybody.

Making sense of the state of privacy in the world of cyber-security systems requires bridging the gap between the cyber-security literature and the privacy literature. Several articles survey cyber-security threats and solutions in diverse areas such as distributed systems (Uzunov et al. 2012; Uzunov and Fernandez 2014), cloud computing (Fernandes et al. 2014; Modi et al. 2013a, 2013b), wireless networks (Butun et al. 2014), smart grids (Liu et al. 2012; Yan et al. 2012), Internet of Things (IoT) (Weber 2010), mobile computing (La Polla et al. 2013), collaborative intrusion detection (Zhou et al. 2010a; Vasilomanolakis et al. 2015a), and, more generally, in any type of computing area (Jang-Jaccard and Nepal 2014; Uzunov et al. 2015). There are also several existing surveys on privacy risks in various domains, such as smartphones in the workplace (Miller et al. 2012), RFID chips (Weis et al. 2004), health applications (de los Angeles Cosio Leon et al. 2009), cloud computing (Zhou et al. 2010b), personalization systems (Toch et al. 2012), the Internet of Things (Ziegeldorf et al. 2014), and pervasive systems (Bettini and Riboni 2015). However, to the best of our knowledge, there is no systematic analysis of the privacy properties of cyber-security technologies. This type of wide-ranging analysis is necessary to guide future research toward privacy-preserving cyber-security technologies and to assess the privacy risks of existing technologies. In addition to

[1] The data breaches of encryption companies such as RSA (Labs 2011) and DocuSign (Krebs 2017) exemplify this risk.

[2] See, for example, the recent court order (Authority 2016) of the Italian data protection authority against an Italian university that was continuously collecting data associated with MAC addresses, claiming they were required for security purposes (among other purposes).

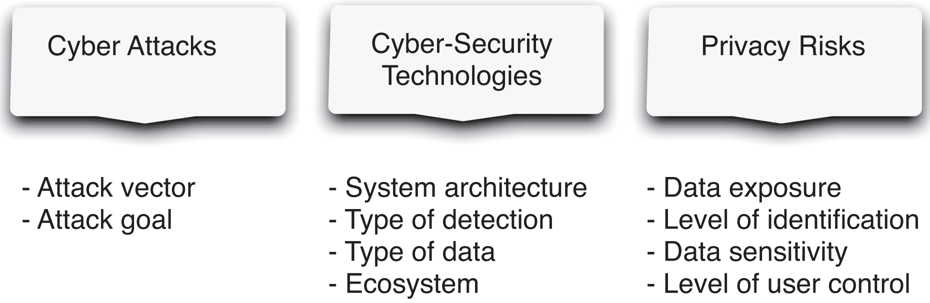

Fig. 1. A visualization of the taxonomy used to analyze cyber-security mechanisms and the risks they pose.

being an important application of online monitoring, cyber-security entertains a balance between external threats, such as the hackers the surveillance can defend against, and the internal threats that are caused by the surveillance itself (Crossler and Bélanger 2017). Therefore, a privacy impact assessment process that is applied to cyber-security technologies in a comprehensive way should take into account both the protective and the intrusive characteristics of these technologies.

The objective of this article is to provide a comprehensive mapping of the privacy threats related to cyber-security technologies. Figure 1 provides an overview of the taxonomy, which details the privacy risks related to different types of cyber-security technologies, which in turn are related to specific cyber attacks. The rest of the article is structured as follows: we start by providing an overview of the relevant technologies, suggesting a basic categorization of cyber attacks and cyber-security systems (Section 2). We then suggest a classification that categorizes the potential privacy threats (Section 3). We analyze the current state of privacy of cyber-security technologies and demonstrate how our taxonomy can be applied to a representative set of technologies (Section 4). We provide a case study of an IoT cyber-system to demonstrate how our framework can be applied to analyze specific system designs (Section 5). Finally, we discuss the effects of our findings on both policymakers and cyber-security developers (Section 6). We especially emphasize future research directions for developing privacy-enhancing cyber-security mechanisms and identifying possible ways to efficiently balance security and privacy.

2 CYBER-SECURITY MECHANISMS

The aim of cyber-security is to protect networks, computers, programs, and data from attacks and unauthorized access. This section first introduces cyber attacks and provides the language for describing cyber attacks and cyber-security systems. The second part of the section proposes a categorization of cyber-security mechanisms that will be helpful when considering their impact on privacy.

2.1 Classification of Cyber Attacks

A first dimension for classifying an attack is the goal of the attack. This is often related to the way an adversary monetizes the attack (e.g., by stealing information and selling it to advertisers or criminals). Overall, the attack goals fall into one of the following categories (Lala and Panda 2001): (1) *stealing information*, such as data on a device, media files, and user credentials; this action is usually performed by spyware malware; (2) *tracking user information*, i.e., monitoring users' sensitive data (e.g., locations, activities, or health-related data); this action is usually achieved using mobile malware; (3) *taking control of a system*, as is done by Trojan, botnet, and rootkit (Graziano et al. 2016).

A second dimension for classifying an attack is the attack vector and it represents the vulnerability exploited by an adversary to gain access to a network or computer system to perform malicious actions. Attack vectors can be identified at three different layers: (1) *hardware*, (2) *network*, and (3) *application*.

2.1.1 Hardware Attacks. At the hardware level, we find attacks that include manufacturing backdoors, gaining access to memory, and hardware tampering. The common goal of these attacks is twofold: modifying the hardware to access sensitive information and creating a backdoor (Tehranipoor and Koushanfar 2010) (e.g., install an invisible program in the hardware circuit) that can be used to regain access to the compromised machine. Such hardware attacks can be applied to several types of devices, such as network appliances, surveillance systems, and industrial control systems.

2.1.2 Network Attacks. Network attacks can target the network protocol or the network device software, and their goal is either the denial of service or hijacking a network connection to steal sensitive data. Specifically, frequent attacks using vectors at the network layer are Denial of Service (DoS) (Schweitzer et al. 2016), IP spoofing (Thang and Nguyen 2016), and man in the middle attacks (Desmedt 2011).

2.1.3 Application Attacks. At the application level, phishing and client-side web attacks are the most common attack vectors, according to the main security market players (e.g., Symantec (2015b)). These attacks target applications such as e-mail services and browsers, since they are the most exposed to the Internet. Regarding attacks through email, phishing is a form of fraud in which the attacker tries to gather sensitive information, such as credentials and credit card numbers by impersonating a reputable entity or person via email, IM or other communication channels (Fette et al. 2007; Ma et al. 2009). Many application-level attacks make use of social engineering techniques that use humans to compromise systems, manipulating them into carrying the attack through deceit (Krombholz et al. 2015). A common example of client-side web attacks is *Cross-Site Scripting* (XSS), which consists of injecting client-side script code (e.g., JavaScript) into web pages. Such injected code could be used for different purposes, such as to bypass access control or to force a user to execute some actions on a remote website on behalf of the attacker.

A large number of application level attacks can be categorized as *malware* (Lanzi et al. 2010). Malware is any malicious software that an attacker manages to run on the target computer. It is used to gather sensitive information, to gain access to private computer systems, or to perform massive attacks. Malware is defined by its malicious intent, acting contrary to user requirements. Malware can be classified into several categories depending on the design goal. The most common malware categories are mobile malware, botnets, spyware (which transmit personal communications), ransomware (which encrypt a victim's data and force victims to pay to decrypt it), and banking malware (Symantec 2015b). Different techniques are used to install malware on a target system. For example, mobile malware is usually installed via SMS, via unofficial application repositories, or by exploiting vulnerabilities of the OS. Once the malware is installed, it can perform several malicious actions, such as stealing information (in this case, it is also called spyware) or tracking user actions.

2.2 Classifying Cyber-Security Technologies

In this subsection, we propose a classification of cyber-security technologies. It is important to note that commercial products do not necessarily have a direct mapping in our classification system, since they often package different protection mechanisms under the same name (e.g., anti-virus), which are possibly offered both as standalone and client-server architectures and for different ecosystems. In the following subsections, we describe four classification categories, and in Section 2.2.5, we present the classification of well-known cyber-security solutions.

2.2.1 System Architecture. Protection systems are software packages that are designed to be deployed according to a specific architecture. The three main architectures are *standalone*, *centralized client-server*, and *collaborative* architectures.

The first architecture (*standalone*) is an architecture in which the cyber-security mechanism is installed only on the local machine to be protected. Such a configuration can be found in the first generation of anti-virus products (Cristalli et al. 2016), where the system performs the entire detection task on the local machine without passing data across the network.

The second architecture (*centralized client-server*) is composed of a client, which is usually installed on the system to be protected, and a centralized server that runs the detection algorithm. This architecture is often adopted by contemporary anti-virus systems when, for example, it has to check whether some visited web domains are malicious or not. The client sends the URL of a particular machine, and the server replies based on its blacklist.

The last architecture (*collaborative*) is implemented as a distributed system, possibly following a peer-to-peer paradigm. It is often adopted by network detection systems such as Snort (Roesch et al. 1999), where sensors are localized on different network nodes and cooperate with each other using a correlation algorithm to determine anomalies/attacks on the monitored network. Recent examples include Worminator, a collaborative intrusion detection system based on encoding threats using Bloom filters (Locasto et al. 2005; Vasilomanolakis et al. 2015a), and other works based on hidden Markov random field (Xie et al. 2016) and autonomic and self-organizing hive-like collaboration (Korczynski et al. 2016).

2.2.2 Type of Detection. Defense mechanisms can operate at the same three levels defined for the attack model (hardware, network, and application) and can be broadly classified in two main categories: *anomaly-based detection*, which learns the routine behavior of a user or application and tries to capture anomalies, i.e., the deviations from the routine behavior (Garcia-Teodoro et al. 2009; Continella et al. 2017), and *signature-based detection*, which tries to characterize the generic behavior of an attack as a signature and then monitors the system, detecting an attack when the signature is observed. There are two main approaches for the last category: the automatic approach builds the signature by using behavioral analysis (system calls, function calls, etc.), while the manual approach requires security experts to explicitly construct the signature by specifying the malicious behavior (Cannady 1998).

2.2.3 Type of Data. Security systems can also be distinguished based on the type of data that their detection algorithm processes. For example, a network intrusion detection system such as Snort (Roesch et al. 1999) analyzes network packets at different network protocol levels, while a host intrusion detection system analyses system call operations performed by an application running on a host. We classify the data used by security technologies into three main categories: *(a) application data*, *(b) file data*, and *(c) network data.* The first category includes both system calls performed by applications and application level data exchanged on the network. For system calls and function libraries, some mechanisms look only at the call itself, while others also inspect the specific parameters of the call; similarly for HTTP protocol requests or emails, some mechanisms look only at the header (e.g., for HTTP, they look only at the request line, i.e., GET and POST commands), while others also inspect the body (e.g., the data being posted with an HTTP request).

In the second category (*file data*), we consider the files that are inspected to ensure that they do not hide a security threat. The most relevant ones are Microsoft Office documents, PDF documents, media files (video, pictures, etc.), and executable files. In the third category (*network data*), we have information contained in low-level network packets. Technically, low level refers to levels below application. Different mechanisms may inspect both the packet header and the contained data or the header only.

2.2.4 Ecosystems. Another dimension that we use to classify security systems is associated with the ecosystems in which a detection mechanism can be applied. In particular, we can apply defensive mechanisms in three main ecosystems: *enterprise*, *mobile devices*, and *IoT*. The enterprise ecosystem represents the typical organization infrastructure, which is composed of locally connected PCs, servers, and network devices, but can also be extended to the use of private and public cloud and web technologies. The mobile devices ecosystem is composed of personal devices typically used in mobility (e.g., smartphones and tablets). The Internet of Things ecosystem is just emerging, but it is already posing serious security concerns. It includes IP-enabled devices (e.g., netcams and smart appliances) as well as sensor networks synchronized to IP-enabled hubs.

2.2.5 Mapping Well-Known Cyber-Security Solutions to Our Classification. Table 1 summarizes several cyber-security solutions according to the classification previously described. We differentiate the solutions according to the source of the analyzed data: (a) network solutions include organization firewalls and Network Intrusion Detection Systems (NIDS); (b) content filtering solutions include proxies, web client-side attack detection, and email phishing and spam detection; (c) endpoint solutions include Host-based Intrusion Detection Systems (HIDS) and Host-based Intrusion Prevention Systems (HIPS) that usually monitor a device (system calls, file system integrity, etc.) detecting malware; finally, we list the general category of security suites, which include commercial products that typically offer a combination of technologies listed under the previous categories. For example, anti-virus products now commonly include prevention of web scripting and phishing attacks. In this category, we also find all-round security solutions that come as a black-box (sometimes called *security appliances* or *next generation firewalls*), as offered by companies such as PaloAlto Networks (Paloalto 2017) or Checkpoint (2017). Note that the term *endpoint solutions* refers to the fact that host data processing is monitored, but it does not imply that the tools implementing the technologies run in standalone mode on the host. For example, some anti-virus tools are composed of two software applications, one running on the server side performing the detection and the other one on the client collecting system information and sending them to the server component. This architecture is mostly used in mobile environments where there are computation and energy constraints. For each category, we provide references to commercial products or scientific articles describing specific systems or techniques and classify them according to the proposed dimensions. Note that certain systems, e.g., Snort (Roesch et al. 1999) (a well known network intrusion detection system), appear in several cells of our table. This is because Snort, as other tools, can be configured in different ways that match different categories.

3 A MODEL OF PRIVACY RISK ASSESSMENT

In this section, we suggest a general methodology for evaluating the impact that a cyber-security technology has on the privacy of the people being monitored. We base our definition of privacy mainly on the theory of contextual integrity (Nissenbaum 2004). We say that a certain technology threatens privacy when private information is accessed in a way that can be used against the original information norms and the control of the individual. For a threat to materialize, an adversary should have the ability to associate a user's identity with data that is considered private by the individual (Solove 2006). As a consequence, to evaluate the privacy threat, we need to carefully understand which parts of released data may lead to the inference of private data and which parts can reveal the identity of a user (possibly joined with external information). The first are often called *sensitive attributes*, while the second, *quasi-identifiers*.

Table 1. A Classification of Cyber-security Solutions

Cyber Protection	**Type**	**Network solutions**	**Content filtering**	**Endpoint solutions**	**Security Suites (Anti-virus, Appliances, . . .)**
Architecture System	**Client-server**	(Roesch et al. 1999; Paxson 1999)	(Smadi et al. 2015; Nelms et al. 2016)	(Portokalidis et al. 2010; Symantec 2016b)	(Cheng et al. 2007; Symantec 2016b)
	Standalone	(Roesch et al. 1999) (Paxson 1999)	(Vigna et al. 2003) (Kruegel and Vigna 2003) (Fette et al. 2007)	(Lanzi et al. 2010; Canali et al. 2012; Feng et al. 2003; Burguera et al. 2011; Fattori et al. 2015)	(Symantec 2016c)
	Collaborative	(Roesch et al. 1999; Ioannidis et al. 2000; Locasto et al. 2005; Xie et al. 2016; Korczynski et al. 2016)	N/A	N/A	(Cheng et al. 2007)
Detection type	**Anomaly-based**	(Roesch et al. 1999; Paxson 1999; Wang and Stolfo 2004; Li et al. 2013)	(Kruegel and Vigna 2003; Michelakis et al. 2004)	(Hoglund et al. 2000; Portokalidis et al. 2010)	(Cheng et al. 2007; Symantec 2016c)
	Signature-based	(Roesch et al. 1999; Paxson 1999; Danda and Hota 2016)	(Fette et al. 2007; Vigna et al. 2003; Kruegel and Vigna 2003)	(Lanzi et al. 2010; Canali et al. 2012; Feng et al. 2003; Burguera et al. 2011; Tripwire 2017)	(Symantec 2016b)
Ecosystem	**Mobile Devices**	(Enck et al. 2014; Portokalidis et al. 2010)	(Rastogi et al. 2016)	(Portokalidis et al. 2010; Burguera et al. 2011; Qualcomm 2017)	(Cheng et al. 2007)
	IoT	(Symantec 2016b; Danda and Hota 2016)	(Symantec 2016b)	(Symantec 2016b; Danda and Hota 2016)	(Symantec 2016b)
	Enterprise	(Roesch et al. 1999; Paxson 1999; Li et al. 2013)	(Fette et al. 2007; Vigna et al. 2003; Kruegel and Vigna 2003)	(Symantec 2016c)	(Symantec 2016c; Paloalto 2017; Checkpoint 2017)
Type of Data	**Application**	(Roesch et al. 1999; Paloalto 2017; Checkpoint 2017)	(Vigna et al. 2003) (Kruegel and Vigna 2003; Fette et al. 2007; Smadi et al. 2015; Nelms et al. 2016)	(Lanzi et al. 2010; Canali et al. 2012; Feng et al. 2003; Portokalidis et al. 2010; Burguera et al. 2011)	(Cheng et al. 2007; Symantec 2016b, 2016c)
	Files	(Paloalto 2017; Checkpoint 2017)	(Paloalto 2017; Checkpoint 2017)	(Symantec 2016b, 2016c)	(Symantec 2016b, 2016c)
	Network	(Roesch et al. 1999; Ioannidis et al. 2000; Paxson 1999) (Li et al. 2013)	N/A	N/A	(Cheng et al. 2007; Symantec 2016c)

The suggested taxonomy was inspired by several existing methodologies for privacy assessment and analysis. First, we adopt data exposure, data sensitivity, and level of user control concepts from Privacy Impact Assessment (PIA) (Wright and De Hert 2011; Oetzel and Spiekermann 2014) and Surveillance Impact Assessment (SIA) (Wright and Raab 2012). These methodologies are used in practice (Wright 2012; Wadhwa et al. 2015) for the design and deployment of specific information systems to evaluate the risks versus the benefits. PIA processes ensure conformance with legal and regulatory requirements, such as the ones required by the U.S. Department of Homeland Security (Clarke 2009) and the new European Union General Data Protection Regulation (GDPR) (European Union 2016). PIA provides tools to determine the risks and effects of a system's data flows and to evaluate protections and privacy-by-design processes to mitigate potential risks. As we analyze cyber-security technological frameworks rather than specific deployments and projects, we disregard some concepts that cannot be addressed in this abstraction level, such as internal procedures for data protection and other policy-related aspects of deploying the technology. For the same reason, we incorporate elements that are based on well-established privacy fair information practices, such as data exposure that is closely related to architectural choices (Spiekermann and Cranor 2009), anonymity (Samarati 2001), and attribute disclosure (Machanavajjhala et al. 2006), which are further described below.

Our taxonomy includes four main categories that model the privacy assessment of each technology: *data exposure*, *level of identification*, *data sensitivity*, and *level of user control*. In the following subsections, we define, expand, and establish each category within its respective theoretical background. Another important category of privacy assessment is the frequency of data release. In many scenarios, information about users is collected periodically or continuously, e.g., when new information becomes available (e.g., updated location of a user) or when partial views of the data are collected by different parts of an information system (i.e., multiple releases). Consequently, multiple privacy notions have been proposed to model this privacy risk (Wang and Fung 2006; Xiao and Tao 2007; Shmueli and Tassa 2015; Shmueli et al. 2012; Riboni et al. 2012). In cyber-security, systems monitor data constantly; thus, the common case is that systems access the data in multiple releases.

3.1 Data Exposure

To analyze the privacy risk posed by each technology, we need to understand which entities have access to the data and the context in which data exposure occurs. There are two main factors influencing data exposure: One is related to the system software architecture that defines the data flow. The other is related to how data in transit, data at rest, and data in use are protected. Spiekermann and Cranor define network centricity as the "degree to which a user's system relies on a network infrastructure to provide a service, as well as the degree of control a network operator can exercise over a client's operations" (Spiekermann and Cranor 2009). A higher network centricity means that the data are more accessible and controllable by external entities, such as data collectors, service providers, location servers, and cloud infrastructure operators. Figure 2 depicts typical network centricity models: (1) a standalone topology in which a user has full control over a standalone client; (2) a centralized topology that rests on a centralized server that monitors data or communications; (3) an external client-server third-party provider that monitors a network by using the cloud; (4) a collaborative topology that carries out monitoring by using a decentralized architecture.

Data exposure is not only affected by the entities that process the data and store it but also by the mechanisms used for data protection. For example, a standalone architecture in which personal data are stored and used without any protection (encryption or access control) has a potential risk of exposing data to unauthorized parties. On the other hand, a system with high network centricity

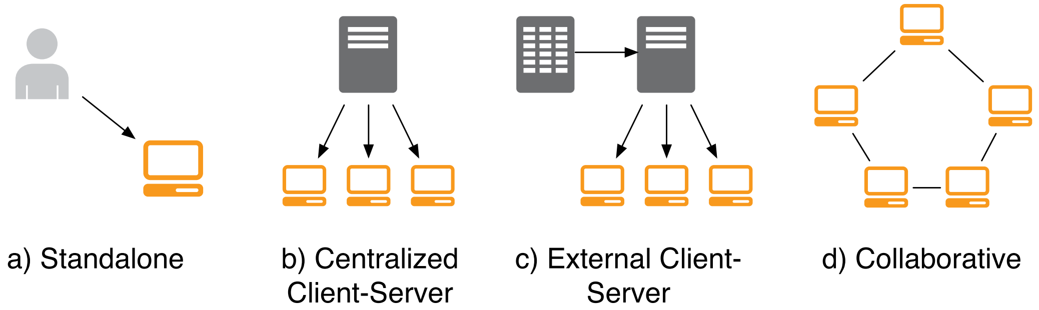

Fig. 2. Examples of network centricity models, each defined by a network topology.

may use secure channels to protect data in transit and encryption and secure computation to protect data at rest and data in use, leading to overall low data exposure (Mascetti et al. 2011).

3.2 Level of Identification

It is well known that the simple removal of explicit identifiers from released data is not sufficient to enforce anonymity (Samarati 2001). Indeed, in several cases, an adversary may re-identify the data respondents by joining part of the released data, called *Quasi-Identifiers* (QIs), with available background knowledge. Consider, for instance, the release of a medical record including personal data such as birth date, gender, and home town. Even if the record does not include the patient's name or social security number (SSN), an adversary having access to a personal information registry may easily limit the set of candidate respondents to those people matching the personal information in the record. In the worst case, the adversary can uniquely re-identify the respondent.

Several privacy notions have been proposed in the literature for measuring the level of re-identification of released data, and algorithms have been devised to enforce those notions in different domains (release of database records, transaction data, statistics, etc.). Perhaps the most known privacy notion in databases is *k-anonymity* (Sweeney 2002): a record is k-anonymous if it can be associated with a set of at least k possible respondents. This privacy notion can be enforced by generalizing QI values such that each record belongs to a group (called a QI-group) of at least k records having identical values for the QI attributes. Assuming that each individual is the respondent of at most one record, each record can be associated with at least k individuals (i.e., the respondents of that record's QI group). Hence, the value k is intended to measure the level of protection from identity disclosure. However, this approach was later shown to have several drawbacks, since the distribution of sensitive values associated with a group of k individuals (even if undistinguishable) has a relevant impact on the privacy risk of revealing the value associated with a specific individual (Li et al. 2007) (e.g., in the case where the same sensitive value is associated with all individuals).

In a different line of work, *unicity* was suggested as a measure of the intrinsic re-identification risk of a dataset (de Montjoye et al. 2013, 2015). Unicity quantifies how much outside information one would need, on average, to re-identify a specific and known user in a simply anonymized dataset. The higher a dataset's unicity is, the higher the probability of successfully re-identifying the user.

The extent to which individuals can be identified is a crucial metric in evaluating the privacy risk. If the tracked data can be easily linked to the real identities of individuals (e.g., to their national identity repositories), then privacy risks have higher chances of being materialized. Unfortunately, in most cases, it is very difficult to realistically model the external knowledge available to an adversary to re-identify individuals. Consequently, properly identifying the set of QI attributes and the level of identification they impose becomes challenging.

In this article, we adopt the three levels of identification suggested by Spiekermann and Cranor (2009), with minor refinement of their definitions:

(1) *Identified*: Unique identifiers, such as ID numbers, names, social security numbers, phone numbers, or other information, can be used to precisely point to a specific individual.
(2) *Re-identifiable* (also called pseudonymous or linkable): Even if the information collected regarding a user does not contain any unique identifier, it can be used to match the user with some other information that is uniquely tied to the individual. Researchers have demonstrated that this is feasible for various data sources, including medical records (Sweeney 2002) and movie ratings (Narayanan and Shmatikov 2008).
(3) *Anonymous*: The user information available is insufficient for re-identification, no matter what inference is performed. When it is shown to be insufficient to limit the set of identities to a set of less than k elements, the term k-anonymous is used (Sweeney 2002).

3.3 Data Sensitivity

As noted above, an adversary can perpetrate a privacy violation from either side of the sensitive association. While k-anonymity provides some level of protection against re-identification, it does not always prevent privacy violations. Indeed, if all the records in a QI-group share the same value for a released sensitive attribute, the adversary can reconstruct the sensitive association between each respondent of the records in that group and the sensitive value. Based on the above weakness of k-anonymity, various privacy notions have been proposed to protect both sides of the sensitive association. In particular, l-diversity (Machanavajjhala et al. 2006) ensures that the records in a QI-group have sufficiently diverse values for the private attribute. Hence, the value l can be used to measure the level of protection from both identity and attribute disclosure. While l-diversity considers *syntactic* diversity among sensitive data, a further notion, named t-closeness (Li et al. 2007), has been introduced to offer additional protection from attribute disclosure based on the semantics of sensitive data.

Anonymity is not enough to ensure privacy in all situations relevant to cyber-security. For example, some services may require authentication with a real identity for billing or for accountability. In these cases, privacy protection techniques need to focus on sensitive attribute disclosure. Obfuscation via generalization is one of the most popular techniques. Generalizing location data, for example, is based on the observation that location information becomes less sensitive when it is less precise, i.e., knowing that an individual is in Manhattan on a given day is usually less sensitive than knowing that an individual is at the address of a particular political party at the time a campaign speech is given (Mascetti et al. 2014).

In this article, we are interested in the different types of sensitive data that can be linked back to a user. Our suggested analysis takes a worst-case approach, aiming to understand which types of data can potentially be monitored by a cyber-security system. Since the types of sensitive data depend on the domain of the technology at hand, we analyze each domain individually, inferring the type of sensitive information that can be obtained from the data. Examples may include web pages visited, email content, and application data. To better understand the sensitivity level of these tracked data items, we turn to the relevant literature. For example, recent studies have shown that many sensitive attributes can be learned using web browsing data, including age and gender (Goel et al. 2012; Hu et al. 2007). Furthermore, as Kosinski et al. show, "Facebook Likes can be used to automatically and accurately predict a range of highly sensitive personal attributes, such as sexual orientation, ethnicity, and religious and political views, personality traits, intelligence" (Kosinski et al. 2013). Similar results were obtained via the analysis of applications used on mobile devices (Staiano et al. 2012; Achara et al. 2015). Network traces (i.e., IP packet traces) can also reveal sensitive information, such as host behavior and network topology (Pang et al. 2006; Ribeiro et al. 2008), even if the network traces are anonymized (Coull et al. 2007).

It is important to note that both sensitive attribute obfuscation and k-anonymity (together with its variants described above) share the same limitation: the promised level of protection is achieved only if the assumptions about the external knowledge available to an adversary hold in practice; if the adversary has additional information, no formal privacy guarantee can be provided.

In the context of privacy-preserving data publication and analysis, other privacy notions have been proposed that do not strictly rely on external knowledge assumptions. In particular, differential privacy (Dwork 2006) guarantees that the probability distribution of query answers to a statistical database is essentially the same, without regard to the existence of a single record in the database. This notion can be enforced by adding noise in a principled way to query answers, and it protects both sides of the sensitive association. Different variants of differential privacy exist. In ϵ-differential privacy, the probabilistic inference of the existence of a single record is bounded by a factor $\exp \epsilon$. In general, values of ϵ less than 0.1 are believed to provide strong protection, while values greater than 10 are considered weak (McSherry and Mahajan 2010).

3.4 Level of User Control

If identified or identifiable personal information is collected, then Fair Information Practices point to mechanisms of user control as policy-based mechanisms that can mitigate privacy threats (Wright and De Hert 2011). Specifically, privacy impact assessment methods describe how notices and choices can inform users about data practices and provide meaningful controls to the user (Oetzel and Spiekermann 2014; of the Australian Information Commissioner 2014). Moreover, for technologies to be perceived as trusted, they are required to provide users with the ability to view a comprehensive or a short privacy policy that includes information about the collection, analysis, use, processing, exposure, and transfer of personal data (Pollach 2007).

However, it is very challenging to design supportive technological interfaces that provide users appropriate ad hoc notices regarding data collection and use choices. According to Spiekermann and Cranor, meaningful and timely information "can be offered with minimal disruption by positioning notices at the point in an interaction where they are most relevant, by providing information in a format that succinctly conveys the most important information, and by limiting notices to situations that are most likely to raise privacy concerns" (Spiekermann and Cranor 2009). Beyond notices, we can measure the level of control that technologies give their users in specifying preferences and have them applied to current and future situations.

To analyze the level of user control, we look at the ability of a technology to interact with a user. While the actual notice and choice relies on the actual implementation of the technology, a preliminary condition for applying them is the underlying interaction model.

We differentiate between three general categories that characterize the possible interactions between a technology and a user:

(1) *No control*: In this category, a system cannot interact with a user in a straightforward manner, making control a hard feature to implement. For example, if a system monitors low-level networking protocols on backbone Internet routers, then interacting with the end user is extremely hard to achieve. As a result, displaying privacy policies or asking the user for her preferences is almost impossible.

(2) *Indirect control*: In this category, a system has the potential to interact with a user; therefore, in certain implementations, the privacy aspect of the technology can be controlled by the user. For example, if the technology is installed on the computer of a user, such as

a local anti-virus system, then the user can uninstall the application, or the designers of the technology can provide notices and choices to the user.

(3) *Full control*: There is an existing infrastructure in which a user can interact with the technology, be notified about privacy and even express her privacy preferences. For example, systems that are deployed as mobile applications on the Android or the iOS platform need to explicitly state the data they want to access and to receive the consent of the user.

4 ANALYSIS

In this section, we analyze the privacy implications of cyber-security monitoring technologies by applying the privacy risk assessment taxonomy (Section 3) to the cyber-security defense systems (Section 2). We see several relations between the characteristics of cyber-security technologies and their underlying impacts on privacy. We analyze each of the dimensions and look at the effects in terms of the privacy risk of each design decision related to a dimension.

4.1 Impact of the System Architecture

The specific software architecture of a system determines, among other things, how data are transferred between different entities and which process each entity should execute. Hence, the architecture is directly related to the potential data exposure to different parties. Standalone mechanisms, which are installed on a user's device (e.g., computer, smartphone, or wearable) and operate only locally, provide access only to the user (Symantec 2016c; Portokalidis et al. 2010) and therefore have low network centricity. Most systems, however, funnel data to a centralized server (Cheng et al. 2007; Symantec 2016b; Roesch et al. 1999; Paxson 1999; Portokalidis et al. 2010).

As explained in Section 3.1, the actual level of data exposure depends also on how data are protected while stored, while in use (processing) and when in transit. While most defense mechanisms protect data in transit, low protection is usually offered for data at rest and in use.

The exposure of unprotected personal data to multiple parties carries a privacy risk. This happens not only in the case of untrusted parties, but also in the case of attacks to these parties, as well as in the case of transfer of part of this data to other parties not explicitly involved in the cyber-security architecture. For instance, governments currently require organizations to share cyber-security information with the government and through collaborative exchanges, information that often includes personal information (Sales 2013).

Many central cyber-security systems profit from sharing information among them regarding cyber threats and sometimes use a 3rd-party provider to enable large-scale collaboration (Vasilomanolakis et al. 2015a). In turn, this can lead to a higher level of network centricity and to an increased threat to privacy. To counter this trend, several recent works relied on a peer-to-peer architecture, in domain areas such as anti-viruses (Cheng et al. 2007), malware detection (Marchetti et al. 2009) and NIDS (Roesch et al. 1999). Peer-to-peer architectures provide the ability to collaborate but with a distributed model in which there is no single owner of the data and where, in principle, the anonymity of users can be better guaranteed by applying specific techniques (Ioannidis et al. 2000). Lincoln et al. (2004) used sanitation methods to remove identifying properties such as IP addresses from collaboratively shared datasets. However, it is unclear whether these methods can withhold a data-linking attack. More robust approaches rely on Bloom filters to code suspicious IP addresses (Gross et al. 2004; Locasto et al. 2005; Bianchi et al. 2008; Vasilomanolakis et al. 2015b). Burkhart et al. (2010) suggested using multiparty and privacy-preserving protocols for event correlation, such as intrusions and DDOS attacks. Shi et al. (2007) used queries over encrypted data to enable auditors to decrypt flows whose attributes fit a specific key provided by a trusted authority. Both encrypted analysis and multiparty computation have the added benefit of ensuring that data can only be used by the designated receivers of the data.

4.2 Impact of the Type of Detection

Anomaly-based cyber-security mechanisms need to build a model of the normal behavior of a system. In principle, this does not imply using personal data, as shown by AccessMiner (Lanzi et al. 2010). However, there are mechanisms that aim at building a detailed model by monitoring the specific operations performed by a user; this may imply inspecting data that can reveal personal information (e.g., Snort (Roesch et al. 1999)). Many anomaly recognition systems build the model once during the learning phase (usually on a server) and then transfer it to the client for monitoring and to report any deviation. Since the normal system behavior may change over time, some anomaly systems also continuously or periodically monitor the operations to keep the model updated using active learning techniques (Moskovitch et al. 2009).

Many cyber-security systems support both anomaly and signature detection (such as Snort (Roesch et al. 1999)). Signature-based detection systems build a model of an attack; this does not usually require access to personal data but rather expert knowledge on the attack. During monitoring, however, the system could require access to user data and activity with the purpose of finding a signature match. In some environments, such as in personal computers, this can be done locally (using a standalone architecture), avoiding the exposure of any personal data to third parties. In others, such as an enterprise cloud or in mobile systems, there are systems that need to search for signature matching remotely on the server (e.g., client-server or collaborative systems). To this end, these systems constantly send user and system operations to the server, causing multiple releases of personal data. In this case, signature-based systems may also pose a privacy threat both in terms of identity and attribute disclosure.

4.3 Impact of the Ecosystem

4.3.1 Mobile Ecosystem. The mobile ecosystem is characterized by devices that usually contain very sensitive personal (and business) data, including contacts, communication patterns, and the whereabouts of a user. Hence, one of the major privacy impacts of cyber-security solutions for this ecosystem is the high sensitivity to attribute disclosure. Similarly, these devices contain information that can be used to easily re-identify a user (e.g., contacts and mobile account information).

4.3.2 IoT Ecosystem. A somewhat similar consideration holds for the IoT ecosystem. Considering that, for example, home automation systems may continuously detect home activities, health-care-related IoT systems may reveal medical conditions, and IoT installations in smart environments may include cameras or other systems that can directly or indirectly identify users performing activities in specific places at given times. These ecosystems are commonly characterized by an architecture that includes one or more gateways under the user's control; they also often rely on cloud data processing. When a cyber-security system performs detection on the gateway, the risk of data exposure is negligible, but whenever processing is moved to the cloud, this is no longer true.

4.3.3 Enterprise Ecosystem. Considering the enterprise ecosystem, particularly the typical organization IT infrastructure, user identity is usually disclosed or easily derived using static IP addresses, along with the infrastructure's DNS association, and MAC addresses uniquely associated with an individual. Hence, whenever data that includes an IP address or a MAC address are released as part of the cyber-security system, the risk of identification is high. Concerning attribute disclosure, in principle, this may include presence and location data, financial data, medical data, and performance indicators that may be considered sensitive. In enterprise ecosystems, there is a tendency to use centralized client-server systems managed within the organization. In this case, data exposure is limited, since personal data are not sent to third parties.

4.3.4 Ecosystem and User Control. Current cyber-security systems offer little, if any, control to the user regarding the way data are inspected; hence, privacy control can be considered nonexistent or very low for all current systems. However, the ecosystem often determines the mode of interaction with the user; hence, cyber-security systems for different ecosystems have different potential impacts on the privacy control they may offer. The mobile environment allows users to receive notices (e.g., via notifications or pop-ups) and enable or disable permissions related to apps and services; hence, this ecosystem has the potential for a good level of privacy control; unfortunately, the control is currently limited to the standard permissions that a system platform offers to apps, which are requested only once, i.e., upon installation. Current anti-virus products for the Android platform explicitly ask for nearly complete access to a system (e.g., AVG 2017 asks for approximately 20 permissions related to messages, phone calls, network communication, and storage).

The IoT ecosystem includes many devices that do not allow any direct interaction with the user; hence, interaction is possible only at the administrative level on the gateway or through an interface with a server-side component when present. Enterprise environments usually include PCs and devices with preinstalled and preconfigured software that offer no privacy control to the end user. These environments often also use NIDS and enterprise malware detection systems. Since these systems operate on routers of the organization network, analyzing data at low protocol levels independent of specific applications, it is quite natural that no control is given to the final user. Standalone systems, such as PCs, allow the user some form of control, including installation consent and activation/deactivation of specific protections; however, these actions are currently not associated with any explanation regarding their privacy implications.

4.4 Impact of Monitored Data

The possibility of anonymization in cyber-security is tightly related to the type of monitored data. Each type of data provides different levels of protection for users, ranging from data that is fully identified to data that can be re-identified but only with some significant effort. It is important to note that we have not found a cyber-security technology that claims that it can provide full anonymization to its users, given the existing re-identification body of knowledge. Several systems, such as phishing detectors, spam detectors, and mobile protection systems, have direct access to straightforward identifying information, such as a user's name, email address, phone number, or email content (sent and received). To protect privacy, these systems require complex mechanisms that can obfuscate identifying details to restore anonymity (Di Castro et al. 2016).

In many cyber-security mechanisms, the monitored data do not directly contain the identity of users but are sufficient for re-identifying users. For example, many types of malware detection systems and anti-virus systems require ongoing access to system calls (including their parameters), which often contain information about usernames, passwords, and other types of information that can be used to identify a user. Web monitoring systems access application-level protocols, such as HTTP and SMTP, all of which regularly contain information that can be used to re-identify users. For instance, HTTP header information can reveal personally identifiable information, such as email addresses typed in a web form (Starov et al. 2016) or the specific browser being used (Nikiforakis et al. 2013). Similarly, mobile protection systems access information that can be used to re-identify users, including the configuration of mobile devices (Kurtz et al. 2016).

Cyber-security systems that operate using network data may be limited to inspecting network packet headers or accessing the content handled by higher-level protocols. In the first case, the IP addresses and logical ports of both senders and receivers are revealed. We have already noted that within organizations, static IPs are often directly mapped to specific individuals. More generally, an IP address can provide information such as approximate geographical location and, combined with other sources, the online services for which an individual has registered, personal interests based on websites visited, and organizational affiliations (Commissioner 2013). The combination of these properties can potentially lead to personal identification.

Considering the privacy risks involved in attribute disclosure, cyber-security technologies that inspect HTTP headers and content in detail may also reconstruct browsing history and the content of query strings, which may be sensitive (and also used as quasi-identifiers (Malandrino and Scarano 2013)). The same holds for email content that, even when not used for re-identification, may be very sensitive. When files are the data being inspected, the file content may be sensitive (relevant for attribute disclosure), but it may also be used for re-identification (e.g., a tagged selfie or biometrics data). The file name itself may sometimes be both sensitive and useful for re-identification. When inspected files are executables, the risk of revealing identifying or sensitive information is much lower. Even when the type of data being inspected is network packets, the data can reveal the sender and receiver hosts in terms of IP addresses. When the data are not limited to the header but also include the content, unless encrypted, the whole communication can be exposed, making the risk of sensitive attribute disclosure potentially high.

4.5 Analysis Summary

The analysis reveals that different aspects of cyber-security systems have different impacts on possible privacy breaches due to the exposure of identifiable personal data. Despite the analysis being quite involved and dependent on the specific design choices of a system with respect to different dimensions, a high-level consideration emerges: The *system architecture* has the most impact on *data exposure*, the *type of data* being inspected influences *identification* and *sensitivity*, and the *ecosystem* impacts *user control* (e.g., mobile devices may easily provide alerts, while IoT devices may not have a direct interface) and usually contains particular data sources that can be integrated with the some information to re-identify users.

It is important to consider that certain characteristics of cyber-security technologies may significantly influence others. For example, anomaly-based detection is usually performed in a centralized client-server architecture, since a server has more computational resources and can improve the accuracy of the model by considering data from multiple systems. Hence, based on our analysis, privacy risks due to correlated cyber-security properties will lead to combined risks.

Table 2 summarizes our considerations and can guide the analysis of a specific cyber-security system, suggesting for each characteristic of the specific cyber-security solution which privacy aspects should be considered to evaluate the related risk.

5 CASE STUDY: INFORMATION SECURITY FOR SMART HOME AUTOMATION

We present a case study of a smart home automation system to illustrate how our guidelines can be applied to the analysis of security systems. The considered use case reproduces the general architecture that is representative of different mainstream products. In particular, we consider the general architecture assumed by a major security market player in a recent technical report (Symantec 2015a). The system is depicted in Figure 3. Several sensors are deployed at home, including presence sensors, power meters, wall switch sensors, smart thermostats, gas sensors, microphones, and cameras. Sensors periodically communicate their data via a wireless network protocol such as Z-Wave to the home gateway. Through the gateway, a remote service provider can update the sensor firmware, send configuration messages to sensors, and send commands to actuators. Moreover, specific gateways may run custom applications that directly control sensors and actuators. Web applications allow the user to configure the home automation directives, to enable or disable sensors, and to send commands to actuators. The user can access these apps either locally or remotely using a mobile device. Authentication is based on passwords. Remote access to the gateway can be

Table 2. Guiding Principles for Privacy Analysis of Cyber-security Systems

Cyber-security system features			**Privacy Implications and suggested analysis**
Dimension	**Type**	**Sub-type**	
Architecture of system	Standalone		[Exp] Low network centricity, but check data at rest and data in use protection.
	Client-server		[Exp] High network centricity. Check policies, data in transit and data at rest/use protection for server and third party entities.
	Collaborative		[Exp] High network centricity. Check data in transit and data at rest/use protection for sharing users.
Type of Detection	Anomaly-based		[Id] [Sens] Re-identifying and sensitive data accessed during training and monitoring. Check specific data and degree of re-identifiability/sensitivity [Freq] Check risk of correlation between multiple releases used to update the model
	Signature-based		[Id] [Sens] Identifiable and sensitive data accessed during monitoring. Check specific data and degree of re-identifiability/sensitivity
Ecosystem	Mobile Devices		[Id] Check for re-identifying data as contacts & location. [Sens] Check for access to communication patterns. [Ctrl] Check level of user control (limited app permissions).
	IoT		[Id] Check for re-identifying data (e.g., video from cameras) [Sens] Check for access to sensible data on medical conditions, behavioral data [Exp] check architecture (Gateway, cloud) and its risks
	Enterprise		[Id] Check for re-identifying data (e.g., static IP/MAC address) [Sens] Check for sensible data as presence, financial records, work performance [Exp] usually medium network centricity (private cloud/server) but check data protection [Ctrl] check for control/notification on PCs. Check for network monitoring without notification.
Type of Data	Application	System calls (no param.)	[Id] Usually preserves anonymity [Sens] check if info on used application is sensitive
		System calls with param.	[Id] Re-identifiable (e.g., by name, address from login forms). Estimate risk. [Sens] Revealed sensible data includes application used, files, input/output, network communication (web browsing)
		HTTP header	[Id] Re-identifiable (e.g., by web browsing analysis, username). Estimate risk. [Sens] Revealed sensible data includes Web browsing history, system information
		HTTP header + content	[Id] Re-identifiable (e.g., by web browsing analysis, username). Estimate risk. [Sens] Revealed sensible data includes Web browsing history, system information, cookies, forms, passwords
		Emails	[Id] Identified (e.g., by signature, name, physical address) or re-identifiable (e.g., email address, affiliation). Estimate risk. [Sens] Revealed sensible data includes email content, sender/receiver
	Files	Documents	[Id] Re-identifiable (e.g., by signature). Estimate risk. [Sens] Revealed sensible data includes personal content, business content
		Executables	[Id] Usually preserves anonymity
		Media	[Id] Re-identifiable (e.g., by tagged photos, voice, biometrics) [Sens] Revealed sensible data may be media content
	Network	IP packet header	[Id] Re-identifiable (e.g., by static IP address in organizations, geo-location, organizational affiliation). Estimate risk. [Sens] Revealed sensible data includes sender/receiver hosts (visited websites)
		IP packet header + content	[Id] Re-identifiable (e.g., through IP, username, email). Estimate risk [Sens] Revealed sensible data includes sender/receiver hosts, communication content

Note: [Exp] Data Exposure; [Id] Identification; [Sens] Sensitivity; [Ctrl] User Control; [Freq] Frequency of Release.

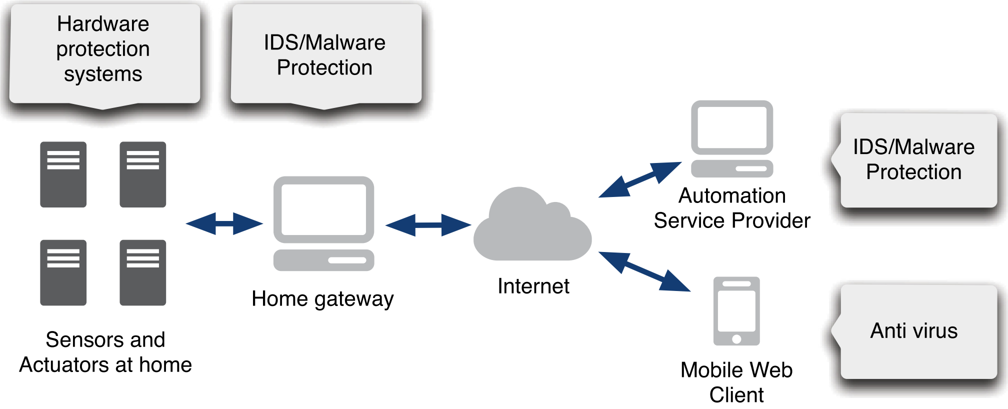

Fig. 3. Case study: A home automation system, with several layers of cyber-security defense systems.

restricted to specific IP addresses. The automation service provider can remotely install, remove, and update apps; it can also update the gateway firmware. Communication between the gateway and external entities is done through the Internet.

As discussed in Symantec (2016a), ONeill et al. (2016), and Lin et al. (2015), an adversary can perform different attacks to achieve different goals. An adversary's goal may be to steal or track personal data (e.g., video streams from cameras, personal habits, and times at which the user is away from home). This goal can be achieved, for instance, by installing malware on a gateway that forwards sensor data to the adversary. Another goal can be to take control of the system to trigger actuator operations (e.g., disabling the anti-theft security system). This goal can be achieved via different methods, including stealing the user's password, installing malware either on the gateway or on the mobile device, or installing malicious firmware on the gateway. Moreover, specific actuator operations can be triggered by tampering with data in transit from sensors to the gateway. The typical attack vectors are the network and application layers.

Unfortunately, a recent study shows that many consumer-level IoT products currently on the market lack effective cyber-security mechanisms, leaving the home automation system exposed to several vulnerabilities (Antonakakis et al. 2017). However, the increasing awareness of users and companies regarding cyber-security, as well as novel data protection regulations, will inevitably determine an improvement of protections mechanisms for IoT systems in the near future. Hence, for the sake of this case study, we assume that the home automation system adopts state-of-the-art cyber-security mechanisms, such as Danda and Hota (2016). In particular, we assume that the system is provided with the following cyber-security mechanisms, which are envisioned in Symantec (2016a) and Sadeghi et al. (2015) as good practice for IoT infrastructures. Communication is protected by encryption and secure authentication. Applications running on a device are protected through code signing. The gateway and the automation service provider infrastructure run HIDS and HIPS for malware detection. The mobile device is provided with a security suite including anti-virus and anti-phishing tools. A tampering detection system monitors the functionality of the sensors.

Next, we analyze the privacy implications of the cyber-security solution according to the guiding principles presented in Section 4. Our analysis is summarized in Table 3.

Architecture of the System. In this case study, the cyber-security defense systems adopt different architectures. The hardware protection system adopts a standalone architecture, while host-based

Table 3. Guiding Principles for Privacy Analysis of the Case Study

Cyber-security system features		Privacy Implications and suggested analysis
Dimension	Type	
Architecture of system	Standalone (hardware protection system), Client-server or Collaborative (host-based security suites)	[Exp] Check whether sensor data are communicated in plain text or not. Communicate sensor data to the gateway through a secure channel (elliptic curve cryptography). Communicate data from the home gateway to remote entities using an encrypted channel.
Type of Detection	Anomaly-based or Signature-based (depending on the used HIDS and HIPS products)	[Id] [Sens] [Freq] Home sensor data, especially if observed on the long term, may reveal sensitive information (presence, routines, medical conditions) or reveal the user's identity. Check specific data and degree of re-identifiability/sensitivity.
Ecosystem	Mobile Devices, IoT	[Sens] High data sensitivity (activities, habits, personal data). Check for sensitive data. Check for risks of inferring additional sensitive data. [Ctrl] Limited user control.
Type of Data	Application, Network	[Sens] Some kinds of sensor data are inherently sensitive (audio-video streams), other kinds may be used to infer sensitive data (number of inhabitants, personal lifestyle. [Id] The user's identity may be reconstructed based on the home gateway IP address or Web activity. Check degree of re-identifiability.

Note: [Exp] Data Exposure; [Id] Identification; [Sens] Sensitivity; [Ctrl] User Control; [Freq] Frequency of Release.

security suites adopt a client-server or a collaborative architecture, depending on the software in use. The network centricity of the hardware protection system is low. However, a recent study conducted on IoT products currently on the market shows that in many cases (19% of evaluated products) the communication of the IoT system with the cloud is not protected by end-to-end encryption (Symantec 2015a). Moreover, in most cases, sensors communicate their data to a gateway using proprietary plain-text protocols to reduce computational and transmission costs, increasing the level of data exposure. Ideally, sensor data should be transmitted using an efficient and secure channel. Technologies such as elliptic curve cryptography may be lightweight enough to fulfill the requirements of several IoT infrastructures. However, the network centricity of HIDS, HIPS, anti-viruses, and anti-phishing tools is relatively high. However, the data exposure is reduced if encrypted channels are used to communicate the data among the home gateway and external entities, and if sensitive data are stored on the home gateway and on the infrastructure of the automation service provider in an encrypted format.

Type of Detection. Depending on the type of detection, different HIDS and HIPS mechanisms may act at different levels. While some technologies may only inspect system calls, other systems may also inspect other information, such as sensor data. For instance, a camera tampering detection system must inspect the video stream of a camera. The same approach may be used to detect tampering with other kinds of sensor data. In particular, anomaly-based detection relies on the long-term observation of a user's data, which may enable the recognition of personal habits and routines and even medical conditions. Moreover, sensor data (e.g., biometric data) and low-level system information (e.g., IP addresses) may lead to the reconstruction of an individual's identity.

Ecosystem. The ecosystem of this case study, which includes IoT and mobile devices, is complex. The complexity of the ecosystem calls for the adoption of different cyber-security mechanisms, which involve a large spectrum of personal data. The ecosystem is characterized by very high data sensitivity, particularly regarding data that involve indoor and outdoor activities, as well as other personal data. Moreover, when matched with external knowledge, this private data may lead to re-identification of an individual.

Type of Data. The cyber-security mechanisms in our use case include an IDS, which is operated by the automation service provider, and may inspect all network data packets. Such packets may include inherently sensitive data such as audio-video streams and other types of data that may be used to infer sensitive information. For example, presence sensor data can reveal the times at which a user is absent from home. Recurrent absence patterns can be inferred by observing the presence sensor data over a long period of time. Sensor-based security systems can have access to hardware-level information, such as fine-grained power meter data. The analysis of such data could yield several kinds of privacy-sensitive information, such as appliances in use in the home, the number of inhabitants, and personal lifestyles (Laughman et al. 2003). Activities carried out in the home can also be recognized based on a combination of different sensor data (Ye et al. 2012). The long-term analysis of recurring activities may enable the recognition of particular kinds of diseases, including cognitive impairment (Hayesa et al. 2008). Even though explicit identifiers are not exposed to the cyber-security mechanisms, other data can be used to re-identify the data owner. The user's identity may be inferred from the (static) IP address of the gateway and from sensor data (e.g., video streams) or from the Web activities observed by the security suite running on the user's mobile device.

User Control. Most users have very limited awareness about what type of data is collected about them using IoT devices, and how this data can be used to infer sensitive information (Egelman et al. 2015). Moreover, most of the IoT devices for smart homes are "closed", i.e., the user has no control on the cyber-security mechanisms actually running on the device. Even in the cases in which some control is offered, lack of awareness can lead to misconfiguration possibly resulting in privacy breaches. The lack of awareness also prevents the users from controlling cyber-security systems operated by the automation service provider, even if those products would offer such a possibility.

6 DISCUSSION

Cyber-security systems pose a unique challenge to privacy. While they protect people's digital safety and information privacy from external threats, our analysis reveals that many cyber-security systems regularly access personal and sensitive information. To balance external and internal threats, we first need to understand the collection of personal information and to evaluate its use against the protection it provides and against other alternatives that maintain adequate security but provide better privacy. In the following section, we discuss the implications of our work on deploying cyber-security systems, developing new privacy enhancing technologies, and regulating cyber-security operations.

Architectural Considerations. Our analysis highlights the complex impact of the type of architecture on the privacy of cyber-security systems. We could not find many instances of contemporary standalone systems, as shared databases are now crucial for recognizing threats and synchronizing responses. It seems reasonable to argue that standalone solutions are not sufficient, though they are considered optimal from the perspective of privacy (Spiekermann and Cranor 2009). Therefore, the design of cyber-security systems should seek other architectures as a means to control access to private information. Collaborative intrusion detection systems offer an interesting model for such an architecture. Since private information is explicitly shared between different parties, a diverse set of Privacy Enhancing Technologies (PETs) has been proposed. We observe a trade-off between trust and privacy when evaluating different architectural solutions. For example, in collaborative architectures, there is no inherent trust regarding the other parties and therefore more restrictive privacy enhancing technologies must be used. It is very conceivable that privacy-enhancing solutions developed for collaborative architectures can also be applied to centralized architectures. More broadly, we argue that cyber-security researchers and developers should assume that every system, regardless of its architecture, should be designed as if the information will be shared with untrusted partners. In this way, the design of a system can protect the privacy of its users by default and can further protect that privacy against future cyber attacks and data leakage.

Protecting Privacy with Technology. There are several important challenges in developing and studying new methods for PETs that can be applied to cyber-security. Many of which have been used in the context of sharing network events and intrusions. A small number of solutions were designed to work in centralized environments, such as email filtering (for spam and malware) in large email servers using anonymization methods (Di Castro et al. 2016). These studies provide an initial set of solutions for privacy-oriented cyber-security technologies. However, they are far from sufficient. We see that PETs for centralized architectures are lacking and that the set of obfuscation and anonymization solutions are still very limited.

Future privacy-aware cyber-security solutions should aim at minimizing data exposure, either by reducing network centricity, for example, by performing a larger part of the analysis at the client side or by providing more comprehensive protection to data during transfer, use and rest using sensitivity analysis (Shu et al. 2015) and cryptographic means such as secure multiparty computations (Liang et al. 2015).

Regulation and Policy. In environments in which privacy is difficult to achieve, regulation and policy become the feasible solutions. Privacy-by-design principles call for an analysis of system design elements and determine whether data handling is legitimate or not by considering the privacy regulations of users' countries and the legal and technical precautions taken by an organization (Oetzel and Spiekermann 2014). Even if an organization concludes that all the personal information is handled legitimately, after performing the analysis that we propose, the reports resulting from the analysis will be precious documentation for accountability purposes in the event of a data breach or dispute. Moreover, a detailed map of personal data processing and related privacy impact assessments are increasingly becoming a standard requirement in several privacy regimes, such as in the European GDPR (European Union 2016).

System designers and regulators have several tools for mitigating privacy threats when architectural solutions and privacy-enhancing technologies do not suffice. The principle of notice and consent is an important fair information practice principles (FIPP), requiring that people should be given notice regarding cyber-security information practices before any of their personal information is collected. However, the implementation of notice and consent should be considered carefully. While most privacy policies and end user license agreements (EULAs) describe what type of data is collected by the cyber-security systems, these agreements are criticized for being too long (McDonald and Cranor 2008) and unreadable by the average Internet user (Sumeeth et al. 2012). It is suggested to use notice and consent mechanisms that provide users with standardized interfaces that allow them to compare and understand the privacy practices of services (Kelley et al. 2009), with timely and contextualized information when deciding which service to use (Schaub et al. 2015), and with choices that are actionable and are tailored to users' preferences (Hirschprung et al. 2017).

Legal experts may rely on security experts or software engineers to evaluate technical aspects of data management. Unfortunately, it is often the case that these experts do not have specific training in privacy and may oversee important issues. An interesting case study is a court order (Authority 2016) of the Italian data protection authority involving an Italian university that was continuously collecting data associated with MAC addresses for several reasons, including cyber-security. The data protection authority rejected the university's claim that they were handling anonymous data. Therefore, we argue that organizations and developers should focus on privacy engineering principles and should make the distinction between privacy and security clearer.

Balancing Privacy and Safety. Embedding privacy in cyber-security requires a new type of thinking regarding the balance of security and privacy. In many cases, the enhancement of privacy can degrade the resolution of data and therefore the effectiveness of security system. Analyzing this trade-off is an open problem, as it requires established methods to measure both security outcomes and privacy. For example, Jin et al. (2017) aimed to characterize the trade-off between intrusion detection accuracy and the privacy of organizations in collaborative intrusion detection and showed that a stable equilibrium is possible under several assumptions. An important future work can be to identify and develop new metrics for quantifying the level of security protection and privacy harm of a given system. These types of measures can help guide the future development and regulation of privacy-oriented cyber-security systems.

7 CONCLUSIONS

The taxonomy presented in this article suggests a way to classify and analyze the privacy implications of cyber-security defense systems. We find that almost all cyber-security technological categories require some access to personal sensitive information, under reasonable assumptions of re-identification and computing power. Our analysis reveals that evaluating the privacy risks involved in using a cyber-security system requires more than the system's generic technological category; it is necessary to classify them in terms of the dimensions identified in Section 2 and to carefully consider their impact on privacy risks, as discussed in Section 4.

Since different privacy-preserving techniques have been proposed to mitigate specific privacy threats (e.g., anonymization and obfuscation for decreasing the sensitivity of the personal data), we believe that the identification of the privacy risks involved in a specific aspect of a cyber-security technology can offer guidance not only in choosing one technique over another but, more importantly, in designing more privacy-aware cyber-security technologies with little or no compromise with regard to their effectiveness in protecting from cyber attacks.

For policymakers, this analysis can be used to guide the regulation, checks, and design requirements that follow the development of a technology. This is particularly important against the backdrop of legislation and policies that set the cyber-security requirements of government agencies and companies. This analysis can also serve as a framework for analyzing the trade-off between the risk that cyber-security systems protect against and the privacy risk that is imposed by the systems themselves. Because our analysis shows that users can be identified in nearly all cyber-security systems, we argue that policies should emphasize the embedding of privacy protections and controls in cyber-security requirements.

Our analysis shows the need of designing privacy-enhancing technologies for cyber-security mechanisms. Privacy-aware cyber-security solutions can control data exposure by choosing privacy-oriented architectures (such as client side or distributed) or by providing more advanced protection to monitored data. Such solutions must also be more transparent to users with respect to the type of information they collect, be more explicit about their privacy and security implications, and provide a better level of control.

ACKNOWLEDGMENTS

The authors thank Dr. Hadas Chassidim for her comments on this article. The authorsalso thank the anonymous reviewers for their valuable suggestions, which improve the content and quality of our article.

REFERENCES

Jagdish Prasad Achara, Gergely Acs, and Claude Castelluccia. 2015. On the unicity of smartphone applications. In *Proceedings of the 14th ACM Workshop on Privacy in the Electronic Society*. ACM, 27–36.

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. 2017. Understanding the mirai botnet. In *Proceedings of the 26th USENIX Security Symposium (USENIX Security'17)*. USENIX Association, Vancouver, BC, 1093–1110. Retrieved from https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis.

Italian Data Protection Authority. 2016. Processing of personal data of employees by e-mail and other work tools. Retrieved from http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/5408460.

Claudio Bettini and Daniele Riboni. 2015. Privacy protection in pervasive systems: State of the art and technical challenges. *Pervas. Mobile Comput.* 17 (2015), 159–174. DOI: http://dx.doi.org/10.1016/j.pmcj.2014.09.010

Giuseppe Bianchi, Simone Teofili, and Matteo Pomposini. 2008. New directions in privacy-preserving anomaly detection for network traffic. In *Proceedings of the 1st ACM Workshop on Network Data Anonymization*. ACM, 11–18.

Iker Burguera, Urko Zurutuza, and Simin Nadjm-Tehrani. 2011. Crowdroid: Behavior-based malware detection system for android. In *Proceedings of the 1st ACM Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM'11)*. ACM, 15–26. DOI: http://dx.doi.org/10.1145/2046614.2046619

Martin Burkhart, Mario Strasser, Dilip Many, and Xenofontas Dimitropoulos. 2010. SEPIA: Privacy-preserving aggregation of multi-domain network events and statistics. *Network* 1 (2010), 101101.

Ismail Butun, Salvatore D. Morgera, and Ravi Sankar. 2014. A survey of intrusion detection systems in wireless sensor networks. *IEEE Commun. Surveys Tutor.* 16, 1 (2014), 266–282. DOI: http://dx.doi.org/10.1109/SURV.2013.050113.00191

Davide Canali, Andrea Lanzi, Davide Balzarotti, Christopher Kruegel, Mihai Christodorescu, and Engin Kirda. 2012. A quantitative study of accuracy in system call-based malware detection. In *Proceedings of the 2012 International Symposium on Software Testing and Analysis (ISSTA'12)*. ACM, New York, NY, 122–132. DOI: http://dx.doi.org/10.1145/2338965.2336768

James Cannady. 1998. Artificial neural networks for misuse detection. In *Proceedings of the National Information Systems Security Conference*. 368–81.

Checkpoint. 2017. Checkpoint security appliances. Retrieved from https://www.checkpoint.com/.

Jerry Cheng, Starsky H. Y. Wong, Hao Yang, and Songwu Lu. 2007. SmartSiren: Virus detection and alert for smartphones. In *Proceedings of the 5th International Conference on Mobile Systems, Applications and Services*. 258–271.

Roger Clarke. 2009. Privacy impact assessment: Its origins and development. *Comput. Law Secur. Rev.* 25, 2 (2009), 123–135.

Canada Privacy Commissioner. 2013. *What an IP Address Can Reveal About You*. Technical Report. Office of the Privacy Commissioner of Canada.

Andrea Continella, Michele Carminati, Mario Polino, Andrea Lanzi, Stefano Zanero, and Federico Maggi. 2017. Prometheus: Analyzing webinject-based information stealers. *J. Comput. Secur.* Preprint (2017), 1–21.

Scott E. Coull, Charles V. Wright, Fabian Monrose, Michael P. Collins, Michael K. Reiter et al. 2007. Playing devil's advocate: Inferring sensitive information from anonymized network traces. In *Proceedings of the Network and Distributed System Security Symposium (NDSS'07)*, Vol. 7. 35–47.

Stefano Cristalli, Mattia Pagnozzi, Mariano Graziano, Andrea Lanzi, and Davide Balzarotti. 2016. Micro-virtualization memory tracing to detect and prevent spraying attacks. In *Proceedings of the 25th USENIX Security Symposium*

(USENIX Security'16). USENIX Association, Austin, TX, 431–446. Retrieved from https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/stefano.

Robert E. Crossler and France Bélanger. 2017. The mobile privacy-security knowledge gap model: Understanding behaviors. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.

Jagan Mohan Reddy Danda and Chittaranjan Hota. 2016. Attack identification framework for IoT devices. In *Information Systems Design and Intelligent Applications*. Springer, 505–513.

M. de los Angeles Cosio Leon, Juan Ivan Nieto Hipolito, and Jesús Luna García. 2009. A security and privacy survey for WSN in e-health applications. In *Proceedings of the Electronics, Robotics and Automotive Mechanics Conference (CERMA'09)*. IEEE, 125–130.

Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel. 2013. Unique in the crowd: The privacy bounds of human mobility. *Sci. Rep.* 3 (2013).

Yves-Alexandre de Montjoye, Laura Radaelli, Vivek Kumar Singh et al. 2015. Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science* 347, 6221 (2015), 536–539.

Yvo Desmedt. 2011. Man-in-the-middle attack. In *Encyclopedia of Cryptography and Security*. Springer, 759–759.

Dotan Di Castro, Liane Lewin-Eytan, Yoelle Maarek, Ran Wolff, and Eyal Zohar. 2016. Enforcing k-anonymity in web mail auditing. In *Proceedings of the 9th ACM International Conference on Web Search and Data Mining*. ACM, 327–336.

Cynthia Dwork. 2006. Differential privacy. In *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP'06) (Lecture Notes in Computer Science)*, Vol. 4052. Springer, 1–12.

Serge Egelman, Raghudeep Kannavara, and Richard Chow. 2015. Is this thing on?: Crowdsourcing privacy indicators for ubiquitous sensing platforms. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1669–1678.

William Enck, Peter Gilbert, Seungyeop Han, Vasant Tendulkar, Byung-Gon Chun, Landon P. Cox, Jaeyeon Jung, Patrick McDaniel, and Anmol N. Sheth. 2014. TaintDroid: An information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Trans. Comput. Syst. (TOCS)* 32, 2 (2014), 5.

Aristide Fattori, Andrea Lanzi, Davide Balzarotti, and Engin Kirda. 2015. Hypervisor-based malware protection with accessminer. *Comput. Secur.* 52 (2015), 33–50.

Henry Hanping Feng, Oleg M. Kolesnikov, Prahlad Fogla, Wenke Lee, and Weibo Gong. 2003. Anomaly detection using call stack information. In *Proceedings of the 2003 IEEE Symposium on Security and Privacy (SP'03)*. IEEE Computer Society, Washington, DC, 62–75.

Diogo A. B. Fernandes, Liliana F. B. Soares, João V. Gomes, Mário M. Freire, and Pedro R. M. Inácio. 2014. Security issues in cloud environments: A survey. *Int. J. Info. Secur.* 13, 2 (2014), 113–170. DOI: http://dx.doi.org/10.1007/s10207-013-0208-7

Ian Fette, Norman Sadeh, and Anthony Tomasic. 2007. Learning to detect phishing emails. In *Proceedings of the 16th International Conference on World Wide Web*. ACM, 649–656.

Pedro Garcia-Teodoro, J Diaz-Verdejo, Gabriel Maciá-Fernández, and Enrique Vázquez. 2009. Anomaly-based network intrusion detection: Techniques, systems and challenges. *Comput. Secur.* 28, 1 (2009), 18–28.

Sharad Goel, J. M. Hofman, and M. Irmak Sirer. 2012. Who does what on the web: Studying web browsing behavior at scale. In *Proceedings of the International Conference on Weblogs and Social Media*. 130–137.

Mariano Graziano, Lorenzo Flore, Andrea Lanzi, and Davide Balzarotti. 2016. Subverting operating system properties through evolutionary DKOM attacks. In *Proceedings of the 13th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, Vol. 9721 (DIMVA'16)*. Springer-Verlag New York, Inc., New York, NY, 3–24. DOI: http://dx.doi.org/10.1007/978-3-319-40667-1_1

Philip Gross, Janak Parekh, and Gail Kaiser. 2004. Secure selecticast for collaborative intrusion detection systems. In *Proceedings of the 3rd International Workshop on Distributed Event-Based Systems (DEBS'04)*. IET.

Tamara L. Hayesa, Francena Abendroth, Andre Adami, Misha Pavel, Tracy A. Zitzelberger, and Jeffrey A. Kaye. 2008. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. *Alzheimers Dement.* 4, 6 (2008), 395–405.

Ron Hirschprung, Eran Toch, Hadas Schwartz-Chassidim, Tamir Mendel, and Oded Maimon. 2017. Analyzing and optimizing access control choice architectures in online social networks. *ACM Trans. Intell. Syst. Technol. (TIST)* 8, 4 (2017), 57.

Albert J. Hoglund, Kimmo Hatonen, and Antti S. Sorvari. 2000. A computer host-based user anomaly detection system using the self-organizing map. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN'00)*, Vol. 5. IEEE, 411–416.

Jian Hu, Hua-Jun Zeng, Hua Li, Cheng Niu, and Zheng Chen. 2007. Demographic prediction based on user's browsing behavior. In *Proceedings of the 16th International Conference on World Wide Web*. ACM, 151–160.

Sotiris Ioannidis, Angelos D. Keromytis, Steve M. Bellovin, and Jonathan M. Smith. 2000. Implementing a distributed firewall. In *Proceedings of the 7th ACM Conference on Computer and Communications Security (CCS'00)*. ACM, 190–199. DOI: http://dx.doi.org/10.1145/352600.353052

Julian Jang-Jaccard and Surya Nepal. 2014. A survey of emerging threats in cybersecurity. *J. Comput. Syst. Sci.* 80, 5 (2014), 973–993. DOI: http://dx.doi.org/10.1016/j.jcss.2014.02.005

Richeng Jin, Xiaofan He, and Huaiyu Dai. 2017. On the tradeoff between privacy and utility in collaborative intrusion detection systems-A game theoretical approach. In *Proceedings of the Hot Topics in Science of Security: Symposium and Bootcamp.* ACM, 45–51.

Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. 2009. A nutrition label for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security.* ACM, 4.

Maciej Korczynski, Ali Hamieh, Jun Ho Huh, Henrik Holm, S. Raj Rajagopalan, and Nina H. Fefferman. 2016. Hive oversight for network intrusion early warning using DIAMoND: A bee-inspired method for fully distributed cyber defense. *IEEE Commun. Mag.* 54, 6 (2016), 60–67.

Michal Kosinski, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proc. Nat. Acad. Sci.* 110, 15 (2013), 5802–5805.

Brian Krebs. 2017. Breach at DocuSign Led to Targeted Email Malware Campaign. Retrieved from https://krebsonsecurity.com/2017/05/breach-at-docusign-led-to-targeted-email-malware-campaign/.

Katharina Krombholz, Heidelinde Hobel, Markus Huber, and Edgar Weippl. 2015. Advanced social engineering attacks. *J. Info. Secur. Appl.* 22 (2015), 113–122.

Christopher Kruegel and Giovanni Vigna. 2003. Anomaly detection of web-based attacks. In *Proceedings of the 10th ACM Conference on Computer and Communications Security (CCS'03).* ACM, New York, NY, 251–261. DOI: http://dx.doi.org/10.1145/948109.948144

Andreas Kurtz, Hugo Gascon, Tobias Becker, Konrad Rieck, and Felix Freiling. 2016. Fingerprinting mobile devices using personalized configurations. *Proc. Privacy Enhanc. Technol.* 2016, 1 (2016), 4–19.

Mariantonietta La Polla, Fabio Martinelli, and Daniele Sgandurra. 2013. A survey on security for mobile devices. *IEEE Commun. Surveys Tutor.* 15, 1 (2013), 446–471.

RSA FraudAction Research Labs. 2011. Anatomy of an attack. Retrieved from http://blogs.rsa.com/anatomy-of-an-attack/.

Chandana Lala and Brajendra Panda. 2001. Evaluating damage from cyber attacks: A model and analysis. *IEEE Trans. Syst., Man, Cybernet.—Part A: Syst. Hum.* 31, 4 (2001), 300–310.

Susan Landau. 2014. Highlights from making sense of snowden, part II: What's significant in the NSA revelations. *IEEE Secur. Priv.* 12, 1 (2014), 62–64.

Andrea Lanzi, Davide Balzarotti, Christopher Kruegel, Mihai Christodorescu, and Engin Kirda. 2010. AccessMiner: Using system-centric models for malware protection. In *Proceedings of the 17th ACM Conference on Computer and Communications Security.* 399–412.

C. Laughman, Kwangduk Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong. 2003. Power signature analysis. *IEEE Power Energy Mag.* 1, 2 (2003), 56–63.

Bingdong Li, Jeff Springer, George Bebis, and Mehmet Hadi Gunes. 2013. A survey of network flow applications. *J. Netw. Comput. Appl.* 36, 2 (2013), 567–581.

Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. 2007. t-closeness: Privacy beyond k-anonymity and l-diversity. In *Proceedings of (ICDE'07).* IEEE Computer Society, 106–115.

Kaitai Liang, Willy Susilo, and Joseph K. Liu. 2015. Privacy-preserving ciphertext multi-sharing control for big data storage. *IEEE Trans. Info. Forensics Secur.* 10, 8 (2015), 1578–1589.

Xi-Jun Lin, Lin Sun, and Haipeng Qu. 2015. Insecurity of an anonymous authentication for privacy- preserving IoT target-driven applications. *Comput. Secur.* 48 (2015), 142–149.

Patrick Lincoln, Phillip A. Porras, and Vitaly Shmatikov. 2004. Privacy-preserving sharing and correlation of security alerts. In *Proceedings of the USENIX Security Symposium.* 239–254.

Jing Liu, Yang Xiao, Senior Member, Shuhui Li, Wei Liang, and C. L. Philip Chen. 2012. Cyber security and privacy issues in smart grids. *IEEE Commun. Surveys Tutor.* 14, 4 (2012), 981–997. DOI: http://dx.doi.org/10.1109/SURV.2011.122111.00145

Michael E. Locasto, Janak J. Parekh, Angelos D. Keromytis, and Salvatore J. Stolfo. 2005. Towards collaborative security and p2p intrusion detection. In *Proceedings from the 6th Annual IEEE SMC Information Assurance Workshop (IAW'05).* IEEE, 333–339.

Justin Ma, Lawrence K. Saul, Stefan Savage, and Geoffrey M. Voelker. 2009. Beyond blacklists: Learning to detect malicious web sites from suspicious URLs. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM, 1245–1254.

Ashwin Machanavajjhala, Johannes Gehrke, Daniel Kifer, and Muthuramakrishnan Venkitasubramaniam. 2006. l-diversity: Privacy beyond k-anonymity. In *Proceedings of International Conference on Data Engineering (ICDE'06).* IEEE Computer Society.

Delfina Malandrino and Vittorio Scarano. 2013. Privacy leakage on the web: Diffusion and countermeasures. *Comput. Netw.* 57, 14 (2013), 2833–2855.

Mirco Marchetti, Michele Messori, and Michele Colajanni. 2009. Peer-to-peer architecture for collaborative intrusion and malware detection on a large scale. In *Information Security.* Springer, 475–490.

Sergio Mascetti, Letizia Bertolaja, and Claudio Bettini. 2014. SafeBox: Adaptable spatio-temporal generalization for location privacy protection. *Trans. Data Priv.* 7, 2 (2014), 131–163.

Sergio Mascetti, Dario Freni, Claudio Bettini, X. Sean Wang, and Sushil Jajodia. 2011. Privacy in geo-social networks: Proximity notification with untrusted service providers and curious buddies. *VLDB J.* 20, 4 (2011), 541–566. arXiv:1007.0408.

Aleecia M. McDonald and Lorrie Faith Cranor. 2008. The cost of reading privacy policies. *ISJLP* 4 (2008), 543.

Frank McSherry and Ratul Mahajan. 2010. Differentially-private network trace analysis. In *Proceedings of SIGCOMM.* ACM, 123–134.

Eirinaios Michelakis, Ion Androutsopoulos, Georgios Paliouras, George Sakkis, and Panagiotis Stamatopoulos. 2004. Filtron: A Learning-Based Anti-Spam Filter. In *Proceedings of the 1st Conference on Email and Anti-spam.*

Keith W. Miller, Jeffrey Voas, and George F. Hurlburt. 2012. BYOD: Security and privacy considerations. *IT Profess.* 5 (2012), 53–55.

Chirag Modi, Dhiren Patel, Bhavesh Borisaniya, Avi Patel, and Muttukrishnan Rajarajan. 2013a. A survey on security issues and solutions at different layers of cloud computing. *J. Supercomput.* 63, 2 (2013), 561–592. DOI: http://dx.doi.org/10.1007/s11227-012-0831-5

Chirag Modi, Dhiren Patel, Bhavesh Borisaniya, Hiren Patel, Avi Patel, and Muttukrishnan Rajarajan. 2013b. A survey of intrusion detection techniques in cloud. *J. Netw. Comput. Appl.* 36, 1 (2013), 42–57. DOI: http://dx.doi.org/10.1016/j.jnca.2012.05.003

Robert Moskovitch, Nir Nissim, and Yuval Elovici. 2009. Malicious code detection using active learning. In *Privacy, Security, and Trust in KDD.* Springer, 74–91.

Arvind Narayanan and Vitaly Shmatikov. 2008. Robust de-anonymization of large sparse datasets. In *Proceedings of the IEEE Symposium on Security and Privacy (SP'08).* IEEE, 111–125.

Terry Nelms, Roberto Perdisci, Manos Antonakakis, and Mustaque Ahamad. 2016. Towards measuring and mitigating social engineering software download attacks. In *Proceedings of the 25th USENIX Security Symposium (USENIX Security'16).* USENIX Association, 773–789.

Nick Nikiforakis, Alexandros Kapravelos, Wouter Joosen, Christopher Kruegel, Frank Piessens, and Giovanni Vigna. 2013. Cookieless monster: Exploring the ecosystem of web-based device fingerprinting. In *Proceedings of the 2013 IEEE Symposium on Security and Privacy (SP'13).* IEEE, 541–555.

Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. Law Rev.* 79 (2004), 119.

Andrew Nolan. 2015. Cybersecurity and information sharing: Legal challenges and solutions. *Andrew Nolan Legislative Attorney CRS* (2015).

Marie Caroline Oetzel and Sarah Spiekermann. 2014. A systematic methodology for privacy impact assessments: A design science approach. *Eur. J. Info. Syst.* 23, 2 (2014), 126–150.

Office of the Australian Information Commissioner. 2014. Guide to undertaking privacy impact assessments. Retrieved from https://www.oaic.gov.au/agencies-and-organisations/guides/guide-to-undertaking-privacy-impact-assessments.

Maire ONeill et al. 2016. Insecurity by design: Todays IoT device security problem. *Engineering* 2, 1 (2016), 48–49.

Paloalto. 2017. Paloalto security platform. Retrieved from https://www.paloalto.com/.

Ruoming Pang, Mark Allman, Vern Paxson, and Jason Lee. 2006. The devil and packet trace anonymization. *ACM SIGCOMM Comput. Commun. Rev.* 36, 1 (2006), 29–38.

Vern Paxson. 1999. Bro: A system for detecting network intruders in real-time. *Comput. Netw.* 31, 23 (1999), 2435–2463.

Shari Lawrence Pfleeger, M. Angela Sasse, and Adrian Furnham. 2014. From weakest link to security hero: Transforming staff security behavior. *J. Homeland Secur. Emerg. Manage.* 11, 4 (2014), 489–510.

Irene Pollach. 2007. What's wrong with online privacy policies?*Commun. ACM* 50, 9 (2007), 103–108.

Georgios Portokalidis, Philip Homburg, Kostas Anagnostakis, and Herbert Bos. 2010. Paranoid android: Versatile protection for smartphones. In *Proceedings of the 26th Annual Computer Security Applications Conference (ACSAC'10).* ACM, New York, NY, 347–356. DOI: http://dx.doi.org/10.1145/1920261.1920313

Qualcomm. 2017. Qualcomm Snapdragon Smart Protect. Retrieved from https://www.qualcomm.com/.

Vaibhav Rastogi, Rui Shao, Yan Chen, Xiang Pan, Shihong Zou, and Ryan Riley. 2016. Are these ads safe: Detecting hidden attacks through the mobile app-web interfaces. In *Proceedings of the Network and Distributed System Security Symposium (NDSS'16).*

Regulation (EU). 2016. Regulation 679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/EC (general data protection regulation). *Offic. J. Eur. Union* L119/59 (May 2016).

Bruno F. Ribeiro, Weifeng Chen, Gerome Miklau, and Donald F. Towsley. 2008. Analyzing privacy in enterprise packet trace anonymization. In *Proceedings of the 15th Network and Distributed Systems Security Symposium (NDSS'08).*

Daniele Riboni, Linda Pareschi, and Claudio Bettini. 2012. JS-reduce: Defending your data from sequential background knowledge attacks. *IEEE Trans. Dependable Sec. Comput.* 9, 3 (2012), 387–400.

Martin Roesch and others. 1999. Snort: Lightweight intrusion detection for networks. In *Proceedings of LISA*, Vol. 99. 229–238.

Ahmad-Reza Sadeghi, Christian Wachsmann, and Michael Waidner. 2015. Security and privacy challenges in industrial internet of things. In *Proceedings of the 52nd ACM/EDAC/IEEE Design Automation Conference (DAC'15). IEEE,*1–6.

Nathan Alexander Sales. 2013. Regulating cyber-security. *Northwest. Univ. Law Rev.* 107, 4 (2013), 1503–1568.

P. Samarati. 2001. Protecting respondents' identities in microdata release. *IEEE Trans. Knowl. Data Eng.* 13, 6 (2001), 1010–1027.

Florian Schaub, Rebecca Balebako, Adam L. Durity, and Lorrie Faith Cranor. 2015. A design space for effective privacy notices. In *Proceedings of the 11th Symposium on Usable Privacy and Security (SOUPS'15).* USENIX Association, 1–17.

Nadav Schweitzer, Ariel Stulman, Asaf Shabtai, and Roy David Margalit. 2016. Mitigating denial of service attacks in OLSR protocol using fictitious nodes. *IEEE Trans. Mobile Comput.* 15, 1 (2016), 163–172.

Elaine Shi, John Bethencourt, T. H. Hubert Chan, Dawn Song, and Adrian Perrig. 2007. Multi-dimensional range query over encrypted data. In *Proceedings of the IEEE Symposium on Security and Privacy (SP'07).* IEEE, 350–364.

Erez Shmueli and Tamir Tassa. 2015. Privacy by diversity in sequential releases of databases. *Info. Sci.* 298 (2015), 344–372.

Erez Shmueli, Tamir Tassa, Raz Wasserstein, Bracha Shapira, and Lior Rokach. 2012. Limiting disclosure of sensitive data in sequential releases of databases. *Info. Sci.* 191 (2012), 98–127.

Xiaokui Shu, Danfeng Yao, and Elisa Bertino. 2015. Privacy-preserving detection of sensitive data exposure. *IEEE Trans. Info. Forensics Secur.* 10, 5 (2015), 1092–1103.

Sami Smadi, Nauman Aslam, Li Zhang, Rafe Alasem, and M. A. Hossain. 2015. Detection of phishing emails using data mining algorithms. In *Proceedings of the 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA'15).* IEEE, 1–8.

Daniel J. Solove. 2006. A taxonomy of privacy. *Univ. Penn. Law Rev.* (2006), 477–564.

Sarah Spiekermann and Lorrie Faith Cranor. 2009. Engineering privacy. *IEEE Trans. Softw. Eng.* 35, 1 (2009), 67–82.

Jacopo Staiano, Bruno Lepri, Nadav Aharony, Fabio Pianesi, Sebe Nicu, and Alex Pentland. 2012. Friends don't lie—Inferring personality traits from social network structure. In *Proceedings of the ACM International Conference on Ubiquitous Computing (Ubicomp'12).* ACM, 321–330.

Oleksii Starov, Phillipa Gill, and Nick Nikiforakis. 2016. Are you sure you want to contact us? Quantifying the leakage of PII via website contact forms. *Proc. Priv. Enhanc. Technol.* 2016, 1 (2016), 20–33.

M. Sumeeth, R. Singh, and J. Miller. 2012. Are online privacy policies readable. *Optim. Info. Secur. Adv. Priv. Assur.: New Technol.* (2012), 91.

Latanya Sweeney. 2002. k-anonymity: A model for protecting privacy. *Int. J. Uncertain., Fuzz. Knowl.-Based Syst.* 10, 05 (2002), 557–570.

Symantec. 2015a. Insecurity in the Internet of Things. Retrieved from http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/insecurity-in-the-internet-of-things.pdf.

Symantec. 2015b. *Threat Report Symantec.* Technical Report. Symantec Inc.Retrieved from http://www.symantec.com/security_response/publications/threatreport.jsp.

Symantec. 2016a. An Internet of Things Reference Architecture. Retrieved from http://wso2.com/whitepapers/a-reference-architecture-for-the-internet-of-things/.

Symantec. 2016b. Device Protection for the Internet of Things. Retrieved from https://www.symantec.com/content/dam/symantec/docs/data-sheets/embedded-security-critical-system-protection-en.pdf.

Symantec. 2016c. EndPoint Protection. Retrieved from https://www.symantec.com/content/dam/symantec/docs/data-sheets/endpoint-protection-en.pdf.

Mohammad Tehranipoor and Farinaz Koushanfar. 2010. A survey of hardware Trojan taxonomy and detection. *IEEE Des. Test Comput.* 27, 1 (2010), 10–25.

Omer Tene. 2014. New harm matrix for cybersecurity surveillance, A. *Colo. Tech. LJ* 12 (2014), 391.

Tran Manh Thang and Van Khanh Nguyen. 2016. Synflood spoof source DDoS attack defence based on packet ID anomaly detection-PIDAD. In *Proceedings of the Conference on Information Science and Applications (ICISA'16).* Springer, 739–751.

Eran Toch, Yang Wang, and Lorrie Faith Cranor. 2012. Personalization and privacy: A survey of privacy risks and remedies in personalization-based systems. *User Model. User-Adapt. Interact.* 22, 1–2 (2012), 203–220.

Tripwire. 2017. File Integrity and Change ManagementFIM. Retrieved from https://www.tripwire.com/.

Anton V. Uzunov, Katrina Falkner, and Eduardo B. Fernandez. 2015. A comprehensive pattern-oriented approach to engineering security methodologies. *Info. Softw. Technol.* 57 (2015), 217–247. DOI: http://dx.doi.org/10.1016/j.infsof.2014.09.001

Anton V. Uzunov and Eduardo B. Fernandez. 2014. An extensible pattern-based library and taxonomy of security threats for distributed systems. *Comput. Stand. Interfaces* 36, 4 (2014), 734–747. DOI: http://dx.doi.org/10.1016/j.csi.2013.12.008

Anton V. Uzunov, Eduardo B. Fernandez, and Katrina Falkner. 2012. Engineering security into distributed systems: A survey of methodologies. *J. Univ. Comput. Sci.* 18, 20 (2012), 2920–3006. DOI: http://dx.doi.org/10.3217/jucs-018-20-2920

Emmanouil Vasilomanolakis, Shankar Karuppayah, Max Mühlhäuser, and Mathias Fischer. 2015a. Taxonomy and survey of collaborative intrusion detection. *ACM Comput. Surveys (CSUR)* 47, 4 (2015), 55.

Emmanouil Vasilomanolakis, Matthias Krügl, Carlos Garcia Cordero, Max Mühlhäuser, and Mathias Fischer. 2015b. SkipMon: A locality-aware collaborative intrusion detection system. In *Proceedings of the IEEE 34th International Performance on Computing and Communications Conference (IPCCC'15)*. IEEE, 1–8.

Giovanni Vigna, William Robertson, Vishal Kher, and Richard A. Kemmerer. 2003. A stateful intrusion detection system for world-wide web servers. In *Proceedings of the 19th Annual Computer Security Applications Conference*. IEEE Computer Society, 34. Retrieved from http://dl.acm.org/citation.cfm?id=956415.956437.

Kush Wadhwa, David Barnard-Wills, and David Wright. 2015. The state of the art in societal impact assessment for security research.*Sci. Public Pol. (SPP)* 42, 3 (2015).

Ke Wang and Benjamin Fung. 2006. Anonymizing sequential releases. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 414–423.

Ke Wang and Salvatore J. Stolfo. 2004. Anomalous payload-based network intrusion detection. In *Proceedings of the International Workshop on Recent Advances in Intrusion Detection*. Springer, 203–222.

Merrill Warkentin and Robert Willison. 2009. Behavioral and policy issues in information systems security: The insider threat. *Eur. J. Info. Syst.* 18, 2 (2009), 101.

Rolf H. Weber. 2010. Internet of things—New security and privacy challenges. *Comput. Law Secur. Rev.* 26, 1 (2010), 23–30. DOI: http://dx.doi.org/10.1016/j.clsr.2009.11.008

Stephen A. Weis, Sanjay E. Sarma, Ronald L. Rivest, and Daniel W. Engels. 2004. Security and privacy aspects of low-cost radio frequency identification systems. In *Security in Pervasive Computing*. Springer, 201–212.

David Wright. 2012. The state of the art in privacy impact assessment. *Comput. Law Secur. Rev.* 28, 1 (2012), 54–61.

David Wright and Paul De Hert. 2011. *Privacy Impact Assessment*. Vol. 6. Springer Science & Business Media.

David Wright and Charles D. Raab. 2012. Constructing a surveillance impact assessment. *Comput. Law Secur. Rev.* 28, 6 (2012), 613–626.

Xiaokui Xiao and Yufei Tao. 2007. m-invariance: Towards privacy preserving re-publication of dynamic datasets. In *Proceedings of International Conference on Management of Data (SIGMOD'07)*. ACM, 689–700.

Yi Xie, Yu Wang, Haitao He, Yang Xiang, Shunzheng Yu, and Xincheng Liu. 2016. A general collaborative framework for modeling and perceiving distributed network behavior. *IEEE/ACM Trans. Network.* 24, 5 (2016), 3162–3176.

Ye Yan, Yi Qian, Hamid Sharif, and David Tipper. 2012. A survey on cyber security for smart grid communications. *IEEE Commun. Surveys Tutor.* 14, 4 (2012), 998–1010. DOI: http://dx.doi.org/10.1109/SURV.2012.010912.00035

Juan Ye, Simon Dobson, and Susan McKeever. 2012. Situation identification techniques in pervasive computing: A review. *Pervas. Mobile Comput.* 8, 1 (2012), 36–66.

Chenfeng Vincent Zhou, Christopher Leckie, and Shanika Karunasekera. 2010a. A survey of coordinated attacks and collaborative intrusion detection. *Comput. Secur.* 1, 29 (2010), 124–140.

Minqi Zhou, Rong Zhang, Wei Xie, Weining Qian, and Aoying Zhou. 2010b. Security and privacy in cloud computing: A survey. In *Proceedings of the 2010 6th International Conference on Semantics Knowledge and Grid (SKG'10)*. IEEE, 105–112.

Jan Henrik Ziegeldorf, Oscar Garcia Morchon, and Klaus Wehrle. 2014. Privacy in the internet of things: Threats and challenges. *Secur. Commun. Networks* 7, 12 (2014), 2728–2742. DOI: http://dx.doi.org/10.1002/sec.795 arxiv:1505.07683

Received August 2016; revised June 2017; accepted December 2017