11-2011

Towards a Privacy Management Framework for Distributed Cybersecurity in the New Data Ecology

Travis D. Breaux
Carnegie Mellon University

Catherine B. Lotrionte
Georgetown University

Published In
Proceedings of the IEEE International Conference on Technologies for Homeland Security (HST), 2011, 6-12.

Towards a Privacy Management Framework for Distributed Cybersecurity in the New Data Ecology

Travis D. Breaux
Institute for Software Research
Carnegie Mellon University
Pittsburgh, Pennsylvania, United States
breaux@cs.cmu.edu

Catherine B. Lotrionte
Government and School of Foreign Service
Georgetown University
Washington, District of Columbia, United Stated
lotrionc@georgetown.edu

***Abstract*—Cyber security increasingly depends on advance notice of emerging threats as individuals, groups or nations attempt to exfiltrate information or disrupt systems and services. Advance notice relies on having access to the right information at the right time. This information includes trace digital evidence, distributed across public and private networks that are governed by various privacy policies, inter-agency agreements, federal and state laws and international treaties. To enable rapid and assured information sharing that protects privacy, the US government needs a means to balance privacy with the need to share. In this paper, we review US laws and policies governing government surveillance and describe key elements for a privacy management framework that seeks to enable government investigations while protecting privacy in a systematic way. The framework aligns existing Federal investigative guidelines for attributing a cyberattack with concerns for automated decision making that arise from the Fourth Amendment "reasonable expectation of privacy" and several fair information practice principles. We discuss technical challenges for those seeking to implement this framework.**

Keywords-cyber security, surveillance, privacy

I. INTRODUCTION

The Internet and pervasive, mobile computing are transforming many facets of daily life, including commerce, health, education, transportation and national security. Over the last few decades, information systems have dramatically evolved to allow individuals to tightly integrate their activities in physical space with new types of private and public data in cyber space. This is especially visible in mobile applications, which allow users to plan routes and schedule appointments, "check-in" at restaurants and other venues, share personal thoughts, images and video, and even make bank deposits remotely by photographing bank checks, all within seconds.

While this transformation offers new social, economic and personal affordances, it also signals a change in how we think about cyber security: the increased integration of cyber-physical systems leads to the availability of new criminal opportunities and new cyber threat indications. Consequently, this change has increased the pressure on the US government to engage in pre-emptive collection activities to improve cyber defense, which forces industry, government and the military into a new ecosystem that challenges traditional lines of separation and historical understandings of the law.

Recently, the Department of Homeland Security (DHS) has taken several steps to improve privacy as it seeks to protect civilian government networks. For example, DHS requires fusion centers, which receive Federal grants, to implement privacy policies that are at least as comprehensive as the Information Sharing Environment Privacy Guidelines. This includes conformance to the Fair Information Practice Principles (FIPPs), which are the basis for the Privacy Act of 1974.

In this paper, we consider the policy and technical challenges to address the integration of cyber security measures with law enforcement authorities in the new data ecology. The remainder of this paper is organized as follows: in Section 2, we review policy and law pertaining to government surveillance; in Section 3, we review relevant definitions of privacy and supporting technology; in Section 4, we discuss technical issues for integrating cyber defense and law enforcement surveillance, with our conclusion in Section 5.

II. LAW ENFORCEMENT SURVEILLANCE

Technology and law enforcement surveillance have evolved over several decades. The basis for personal privacy against government surveillance in the United States is the Fourth Amendment of the US Constitution, which guarantees that the government will conduct only reasonable searches and seizures (including through electronic means). Based on federal court interpretations of this requirement, the presumption is that a search or seizure is reasonable if the government obtains a warrant based on probable cause before conducting the search and seizure. The courts, however, have established a number of exceptions to a warrant requirement for searches and seizures, including:

- **Publicly available information**: Information obtained from "public" spaces, which any party can enter and observe
- **Consent**: Information that an individual voluntarily allows the government to surveill
- **Third-party Rule**: Information disclosed to third parties, who proffer the information

According to the Supreme Court in the *Katz* case, an individual's Fourth Amendment rights are assessed based on whether the individual has a "reasonable expectation of privacy" in the information at issue. Courts have reasoned that, if information is in "public spaces" where anyone (including the government) could view that information, then the individual does not have a

"reasonable expectation of privacy" in that information. With respect to the Internet, public spaces may include blogs, forums, and chat rooms, whether or not they require an unrestricted registration to gain access, i.e., everyone must register, but no one is excluded. Private spaces, which are invitation-only and to which law enforcement cannot obtain an invitation would require a warrant prior to search and seizure. Furthermore, Orin Kerr argues that encryption does not provide a reasonable expectation of privacy, because once the ciphertext is in plain view, the government is authorized to try to access it [12].

For private spaces that have multiple members with legal access to the space, any member may consent to a search and seizure to the parts of the space that are determined to be within his or her lawful control. For example, with physical locations like an apartment with multiple tenants, one of the tenants can consent to a search of the premises but that search will be limited to the parts of the premises that the individual has lawful possession of. This may include the individual's room and a common living area, but not other tenants' rooms.

As new technologies emerge, law enforcement seeks to update existing laws to support their investigations. These updates include new surveillance powers commensurate with increased threats (e.g., after 9/11 new authorities were granted to the federal government based on the reasoning that it was the lack of such authorities that limited the government's ability to detect and prevent terrorist activities). The following laws shaped law enforcement access over the past decades:

- 1968: Wiretap Act authorizes access to wire and oral communications with a court order
- 1986: Pen Register Act authorizes access to telephony communications with a court order, if access is *relevant* to an investigation (probable cause not required).
- 1986: Stored Communications Act restricts access to communications in electronic storage to those with a court order or subpoena
- 1994: Communications Assistance for Law Enforcement Act (CALEA) requires telecommunications carriers and manufactures to provide access to lawful communications intercepts
- 2001: USA-PATRIOT Act extends Pen Register Act to include Internet communications, and authorizes the national security letter (a subpoena), which allows access to meta-data, excluding subject matter content

In addition, there are several other laws that permit or restrict the disclosure of personal information to law enforcement for specific purposes without a court order. This includes the Fair Credit Reporting Act (FCRA), which permits disclosure of specific information (name, former addresses, bank accounts) to law enforcement for counter-intelligence activities,[1] and the Health Insurance Portability and Accountability Act (HIPAA), which permits disclosures of specific information (e.g., name, date of birth, blood type) for identifying suspects or missing persons, with the exception of DNA, dental records and samples of body fluid and tissue.[2]

Offensive cyber intrusions often include some level of premeditation, for example, when individuals train in the art of gaining unauthorized access or disrupting services by exploiting computer vulnerabilities. In addition, ideology or other factors may affect an attacker's motive and opportunity for conducting a cyber attack. These indications can present themselves in "digital traces," when the attacker moves between cyber and physical spaces. In seeking to leverage these traces to pre-empt a cyber attack, one must consider how law enforcement transitions from a "hunch" or "lead" to a full investigation in accordance with the legal and policy guidelines that the government operates under.

The U.S. Federal Bureau of Investigation (FBI) Attorney General's Guidelines for Domestic FBI Operations of 2008, also called the Mukasey guidelines, provide guidance to the FBI in the context of this larger legal environment. These guidelines, which were more recently updated, include collection and retention of information about U.S. citizens:

- Describing an unlawful act and that is incidental to an ongoing investigation
- For preparing *assessments*, which support an authorized purpose but not any factual predication
- For preparing preliminary and full investigations, which have a factual basis.

Network anomalies, such as unusual network port or protocol access, unusual data transfers, unexpected or unusual files, or IP address conflicts may be evidence of an emerging, active or persistent cyber attack. These anomalies may be treated as leads to open an assessment, which can then utilize a broader set of information to analyze and discover evidence of a threat. This may include public information indicating the identity of an attacker, or it may be patterns of behavior observable across other corporate or government networks. However, to protect privacy (of individuals and organizations), access to this information must be commensurate with utility to move an investigation forward and the risk and impact of a potential threat.

III. Definitions of Privacy

Integrating outside intelligence to pre-empt or attribute an emerging threat to particular individuals or nation states improves our ability to protect critical infrastructure. Advance notice enables law enforcement and diplomatic options, as opposed to military options, to mitigate these threats before the threat escalates to cause significant damage or disruption [10]. However, intelligence activities to pre-empt such threats carry the risk of invading the privacy of innocent individuals. In the U.S., the definition of privacy is often traced back to Warren

[1] 15 U.S.C. § 1681u

[2] 45 CFR 164.512(f)(2)

and Brandeis' 1897 Supreme Court statement the "right to be left alone" [21]. Since this early definition, however, legal scholars, philosophers and computer scientists have sought more sophisticated definitions based on technological solutions. We discuss two paradigms that affect how technology implements privacy in theory and in practice.

Privacy is Secrecy. For many, privacy is viewed as *secrecy* or *confidentiality*, which is considered a subset of information security. This viewpoint inspires research in anonymity, and its converse identifiability, to understand the technical limits of secrecy. Sweeny and Machanavajjhala et al. found that individuals can be uniquely identified by a combination of attributes, none of which alone uniquely identify the individual [20, 15]. For example, a person's birth date, gender and zip code, which are commonly exchanged for access to online services, were found to uniquely identify 87% of U.S. residents [20]. Alternatively, differential privacy includes methods to de-identify query results from statistical databases to ensure individuals are not identifiable using such queries [6]. The disadvantage in confidentiality is that losing identifiability equates with a loss in information utility; it is identifiability that enables sophisticated analysis to identify emerging threats. It is argued that anonymity not only threatens security but also has been wrongly equated with online privacy [1].

On the other hand, the question has been raised whether information that is not viewable by humans but only computers (in zeros and ones), would have any Fourth Amendment protections. [13]. Under this assumption, we can define a black box containing personally identifiable information and the only output from the black box is a person's identity and a hypothesis concerning suspicious and potentially criminal activity: the premises that led to concluding the hypothesis remain confidential, stored within the black box and inaccessible to humans. With the hypothesis in hand, it remains the independent task of law enforcement to build a factual predication to an investigation. The precision of the analysis (the ratio of true positives to false positives) determines the utility of this approach. The European Union Data Protection Directive 95/46/EC, Article 15 prohibits basing legal or other significant decisions solely on automated processing. The U.S. Computer Matching and Privacy Protection Act of 1988 requires oversight to use "matching programs" across government agency databases, although, verification of program results is only required if a negative impact to individual public benefits or payments is perceivable.

Privacy is Control. For others, privacy is viewed as the effect of individual control. This viewpoint includes several recommendations: (1) individuals should be aware of how their information is collected, used, transferred and retained; (2) individuals should have the right to consent to (opt-out of) such uses, transfers and retention; (3) individuals should have access to view their information and (4) individuals should be permitted to correct their information, if found inaccurate. These recommendations correspond to the *transparency* and *individual participation* principles in FIPPs. The Federal Trade Commission (FTC) has adopted these principles in their regulatory guidance to commercial businesses.

In a law enforcement context, there are obvious limits to the extent of individual control afforded to suspects during an ongoing investigation. In the absence of individual control, additional restrictions can be implemented that afford U.S. citizens a higher degree of privacy. These include other FIPPs that are related to internal information practices: (a) *purpose specification*, whereby organizations explicitly state the purpose for which information will be used; (b) *collection limitation* or *data minimization*, which means organizations will only collect information for specified purposes and only retain information for as long as is necessary to fulfill those purposes; (c) *use limitation*, which means organizations will only use and transfer information for the purposes for which it was originally collected; and (d) *data quality*, which requires organizations to ensure information is accurate, relevant, timely and complete. DHS have incorporated all of the FIPPs into their routine information practices and even pushed the FIPPs out to state, local and tribal partners.

IV. Framework for Privacy Management

Law enforcement needs an evolvable framework, both to manage privacy in the presence of changing technology and law, and to investigate emerging cyber attacks and mitigate threats using available intelligence. To this end, we present several key elements for an information sharing framework, including: (1) inquiry escalation; (2) escalation validity; (3) collection and use rights; and (4) data provenance and quality. While we are not advocating for any specific approach, we believe that these elements are one of several possible outcomes, based on our analysis of existing policy and technology trends. We begin by presenting a hypothetical scenario intended to challenge the limits of our discussion.

At 9:37 pm, a company that manages or develops critical infrastructure, such as a power company or defense contractor, detects an anomaly on their internal network. The company sends data to law enforcement, including a remote IP address, and description of the activity, such as repeated, failed login attempts or a large transmission of encrypted data. Because the anomaly is not in a category of known cyber attack threat indications (e.g., port scans, malicious code signatures), law enforcement only opens an initial assessment.

The company's role in critical infrastructure means the potential impact from a cyber attack is higher than average, so law enforcement obtains a national security letter to identify websites accessed by the IP address, including Internet Service Provider subscriber's name and address. This subscriber information links the anomaly to data that describes the physical space. Several questions ensue that further collecting additional data.

Who else lives at the address? Because IP addresses identify wired or wireless routers that support multiple,

simultaneous connections from different users, law enforcement purchases credit reporting data from LexusNexus. The data includes the full names and state driver's license numbers linked to a particular address: two individuals are targeted, Alice and Bob.

How to reduce the number of human targets in the assessment? Law enforcement maps the driver's license numbers obtained from LexusNexus to a vehicle license registration database to identify license plate numbers of cars driven by the targets. These license plate numbers are cross-referenced to license plate scans obtained using optical character recognition (OCR) technology at traffic intersections, tollbooths and patrol cars. Bob was last seen exiting the toll way toward home at 9:45 pm, which is 18 minutes after the anomaly was detected originating from his home computer network, so law enforcement shifts their attention to Alice. Bob may later receive focus during the assessment, since the anomaly may have been a timed logic bomb.

Has the IP address been the origin of any accesses to known hosts for criminal activity? The list of accessed websites include web page request header information, such as usernames or aliases used on the website. The websites include public, technical forums that may be accessed using a no-cost web account, so law enforcement creates an account and pulls content from the site: messages to/from the identified alias containing links to known hacking tools, which have been accessed by the IP address. Possession of these tools is not illegal.

What other access points did Alice have access to? Location data is sought without warrant: public websites that record location information, such as Foursquare, Facebook Places, Twitter, and so on, for establishments with wireless networks, such as Universities, cafes, friends' homes. IP addresses are cross-referenced with other anomalies reported by the targeted company or other companies in the same industry or risk threshold.

Has Alice had any contact with known suspects or criminals? Law enforcement returns to Alice's location data and cross checks these locations with those of known suspects and criminals. This check discovers that Alice is observed in the same location as a known suspect in another cyber investigation, however, this does not prove she had contact. By cross-checking Alice's e-mail correspondence (the addresses of senders and recipients but not message content), law enforcement discovers she received e-mails from the suspect.

At this point, the assessment yields several facts: Alice lives at the address where the anomaly appears to have originated based on her IP address. The IP address was observed remotely downloading software tools that could be used to generate the anomaly, and Alice had contact with a suspect in an ongoing investigation. Based on pre-determined criteria, these facts could be sufficient to open a preliminary investigation. Note, however, that it is not conclusive that Alice is the culprit in an emerging crime. Rather, Alice's computer (or Bob's) may have been an intermediate host in producing the anomaly. Motivated by existing policy, technology and this scenario, we now turn to the privacy framework.

A. Automated Inquiry Escalation

Automated inquiry escalation refers to the quantity and quality of evidence necessary for automated use of certain investigative data sources and methods (see Table I, based on the Mukasey guidelines). Each stage of an inquiry is initiated by a logical trigger: assessments are triggered by leads (and misleads), which may yield circumstantial evidence or facts that trigger preliminary investigations or full investigations, respectively. Leads (or hunches) are the least reasonable indicators of suspicion, whereas relevant facts are the most indicative. These triggers yield access to additional information sources as evidence accumulates. Laws, memorandum of understanding and other policies govern what information is available at each stage of an inquiry: e.g., by permitting access without a warrant (assessments and pre-investigations) or by requiring a warrant (full investigations). Because these laws change in response to new technologies and emerging social or political issues, a system for automating inquiry escalation must be periodically updated to grant or revoke access to data. These updates entail codifying regulations and information sources in a common policy language.

TABLE I. INFORMATION ACCESS AND INQUIRY ESCALATION

	Assessments	**Preliminary Investigation**	**Full Investigation**
Triggers	Leads and Misleads	Circumstances	Facts
Information Access Matrix			
Public Info	X	X	X
Government Data	X	X	X
Commercial Serv.	X	X	X
Volunteered Data	X	X	X
Subscriber Info.	X	X	X
Location Info.		X	X
Subpoenas		X	X
Pen Registers		X	X
Undercover Ops.		X	X
Elec. Surveillance			X
Search Warrant			X

Triggers can be expressed using pattern languages, such as regular expressions, which map relevant machine-readable patterns onto datasets. For example, leads may be generated from network anomalies in an intrusion detection system, when a remote computer scans ports on a commercial or government network. In this case, the pattern *premise* consists of matching remote and local IP addresses and network requests directed at port numbers across a specified timeframe; the pattern *conclusion* is a suspected port scan, which is a potential indication of an emerging attack. Patterns can be fairly complex, ambiguous or inconclusive and shared across commercial, non-profit and government computer defense systems. By aggregating multiple patterns together, triggers can use a utility function to determine which pattern combinations are sufficient for escalation.

Under the Data Mining Reporting Act (DMRA), these patterns would not be subject to oversight because of an exception that excludes reporting patterns discovered for cyber security. In the event that all-source attribution is successful, law enforcement can data mine new "patterns" of behavior that correlate with known or convicted criminals for cyber security purposes. A risk to privacy exists, if these patterns were then re-purposed to investigate crimes outside of cyber security, that would otherwise be covered by the DMRA. Because the DMRA covers reporting pattern "discovery," but not pattern application, the re-purposing of patterns would not be subject to congressional oversight.

With each escalation, new data sources and methods become available. In addition to commercial services (e.g., LexusNexus) and government databases, a private company may volunteer access to their employee and consumer data for specific purposes, such as terrorism or national security investigations; this data normally requires a grand jury subpoena or court order. We foresee two profiles for volunteering data: "benefit of the commons," or companies that recognize their services provide critical data to identify suspects, such as airlines or banks; and "benefit of a represented class," or companies in the same industry (banking, defense, energy, transportation) or in the same supply chain who may be more likely to share vulnerabilities or fear of becoming victims to repeat attack strategies applied across their class.

Inquiry automation poses several privacy challenges. These include escalation validity, managing collection and use rights, and data provenance and quality.

B. *Automated Escalation Validity*

When properly implemented, escalation can enhance privacy protection by scaling surveillance with an appropriate degree of suspicion, impact and risk. The utility function that triggers escalation should include several calculations: (1) *precision*, or the ratio of true positives (reasonable suspicions) over the total number of escalations generated by the trigger; (2) the *impact factor*, which is the perceived cost of disruption or denial created by a successful offensive cyber operation; and (3) the *risk factor*, which is the probability that the suspected cyber attack will occur based on historical data. The impact and risk factors will vary based on their sensitivity: do they reflect a single organization or a collective? This variance can amplify or diminish perceptions of impact and risk across different contexts.

Figure 1 illustrates these calculations for three hypothetical patterns: the left, vertical axis describes the precision (blue) and risk factor (red); the right, vertical axis describes the impact factor (green) on a coarse scale of one to ten. Pattern #1 has nearly a 50-50 chance of yielding a true positive, i.e., a matching pattern leads to a reasonable suspicion, and the 50% risk factor indicates that a cyber attack associated with this pattern is fairly common. Patterns #2 and #3 have lower precision and are less likely to yield successful investigations. However, Pattern #2 maps to a low risk, high impact factor (e.g., weapon of mass destruction), whereas Pattern #3 maps to a high risk, low impact factor (wire fraud).

Arguments may be made that, despite low precision, Patterns #2 and #3 should escalate an inquiry due to the high impact and high risk, respectively. In the presence of high impact or risk measures, one should consider precision, or the pattern's ability to successfully detect these events. The product of precision and risk yields the likelihood that the pattern predicts a reasonable suspicion that the cyber attack would occur. For Patterns #1, #2, and #3, these scores are 23%, 3% and 7%, respectively. Despite the high risk (70%) for the cyber attack linked to Pattern #3, there is a low probability of 7% that the pattern could lead to an investigation of a real cyber attack. In other words, Pattern #3 just isn't very good at prediction given the high risk of a real attack. Alternatively, some may argue that the high impact factor 9 for Pattern #2 discounts the low probability (3%) of a real attack.

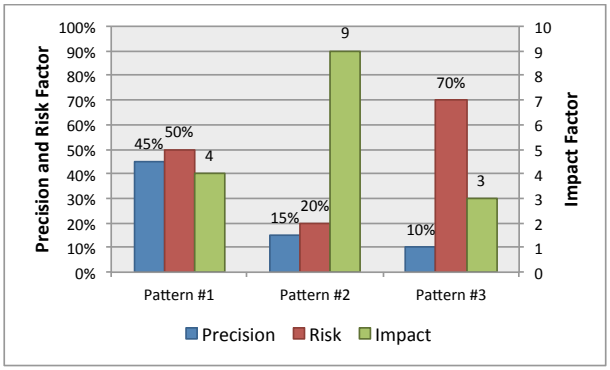

Figure 1. Comparison of three hypothetical patterns for automated escalation validity

These calculations rely on past events and therefore they must be designed into the automated system with high-integrity. With advance knowledge, a malicious actor could target innocent individuals by simply manipulating any two of these calculations. High confidence in these calculations is necessary to address criticism that DHS lacks justification for certain actions [17] and that analysis methods must be subject to science-based assessments [7]. These calculations may provide a predicator for voluntary information sharing, for example, data sources not easily acquired without court orders may be volunteered by companies for certain low risk, high impact purposes (e.g., service outages), but not for high risk, low impact purposes (e.g., minor theft).

In addition, contradictory evidence should be sought in parallel with supporting evidence to increase escalation validity. For example, a fact-finding pattern to de-escalate triggering a full investigation may be based on location information. Location can be used to determine that a suspect could not have accessed a location from which an attack originated, in the event that a suspect's host computer was infected with malicious code to obfuscate the real attack origin. Linking escalation with de-escalation patterns increases protections for innocent

individuals and reduces waste, especially when attackers use deception or obfuscation strategies.

Automation also introduces the risk of repeat escalations that are non-productive. For example, under existing guidelines, assessments can run uninterrupted for 30 days, at which time they are reauthorized or terminated. Unproductive assessments fail to generate circumstantial or fact-based evidence and thus must be terminated. However, a new lead could re-trigger the same assessment and thus an unending cycle could ensue. The fact that leads, which trigger assessments, are based on timely indications is important to permit restarting assessments. However, thresholds or priorities should be established to avoid consuming resources and automation cycles with repeat, non-productive inquiries.

C. Collection and Use Rights

Integrating data sources can violate assumptions about privacy protection when data is shared for specific purposes. Certain laws, such as the HIPAA Privacy Rule and FCRA, permit sharing specific types of information with law enforcement for limited purposes. Commercial, non-profit and government agencies may also voluntarily provide data for limited purposes. The Fair Information Practice Principles (FIPPs), adopted by the DHS Privacy Office, require purpose specification, collection and use limitation as part of a comprehensive privacy program, thus, automated reasoning must operationalize these principles to avoid data leaks and privacy violations.

We propose to use machine-readable, formal languages to specify data flows in a privacy-preserving manner. Formal languages have a formal semantics, which is expressed in first-order logic or another mathematical formalism and which defines the scope of valid interpretations of language expressions. In first-order logic, we can specify a privacy policy and check that this policy complies with a property using *logical entailment*, in which we ask whether a policy P entails a conclusion c, written $P \models c$, if an only if, every interpretation that satisfies the statements in P also satisfies the conclusion [9]. We can use this evaluation to detect contradictions or conflicts in the policy or between policies, which occur when policies do not entail the desired conclusion.

In our policy research [2, 3], we combine Hohfeld's legal concepts, including rights and duties that correspond to "what is permissible" and "what ought to be" in Deontic Logic [8], with notions of classification systems expressed in Description Logic (DL), which is a subset of first-order logic for expressing knowledge in a TBox, or knowledge base [5]. The DL family $\mathcal{ALC}(\mathcal{D})$ provides the ability to reason over rich, real-world descriptions of actors, actions and objects in a policy and what they are permitted to do. The reasoning tasks of deciding if an action complies with a policy using concept satisfiability and concept subsumption in $\mathcal{ALC}(\mathcal{D})$ are PSPACE-complete [14], ensuring tools can scale to larger policies by reason over increasingly larger policies over time.

Consider an example based on DL. We can state that the class of rights R and prohibitions P are disjoint, such that no action in R can co-exist in P; in other words, a right to use data cannot co-exist with a prohibition to use that data. We can then state, with respect to a TBox T, that $T \models (C \setminus D) \sqsubseteq P$ to denote that a data action C to collect e-mail messages, excluding the data action D to collect the message header and routing information, is in the class of prohibitions P; this rule conforms to the rights of FBI assessments shown in Table I. Should we also declare that $T \models C \sqsubseteq R$, we could detect a contradiction in the TBox T, since R and P are disjoint and D is a subset of C. While this example is trivial, it illustrates how DL can be used to detect conflicts in privacy policies across commercial databases and fusion centers or as inquiries escalate to include new data sources. In addition, an extension to this example in $\mathcal{ALC}(\mathcal{D})$ can be used to prove that the FIPPs collection and use limitation principles apply to a set of privacy practices, or the practices do not collect more data than is needed, or use data for purposes beyond which it was originally collected [4]. The following two rules, a right r_0 and a prohibition p_0 expressed in $\mathcal{ALC}(\mathcal{D})$, illustrate in more detail how DL concept subsumption can be used to detect information sharing policy conflicts:

$$r_0 \equiv Permit \sqcap useData.(customerInformation \sqcap \neg customerEmail) \sqcap forIndustry.IndustrialControl$$

$$p_0 \equiv Prohibit \sqcap useData.customerInformation \sqcap forIndustry.(\top \sqcap \neg(Banking)) \sqcap forImpact \leq 4$$

Rule r_0 describes a right to permit using customer information, excluding customer e-mails that are otherwise considered to be a subset of this information, for investigating attacks against the industrial control industry. Alternatively, rule p_0 describes a prohibition to prevent using customer information for any industry (expressed in DL by the symbol $\top$, called "top") except for banking with an impact factor less than or equal to four. A manufacturer might write these two rules as part of their privacy policy to enable information sharing for conducting assessments. The right r_0 permits sharing data to investigate events within the their industrial class or supply chain for events with any impact score (DL uses an open-world assumption). Prohibition p_0 includes an exception for investigating events in the banking industry, because the manufacturer's products are deployed in this sector for example and an exploit in their systems could be used in an attack targeting their customers.

The advantage of expressing positive and negative norms is that policy authors can emphasize their primary goals: share information for a specific purpose, while avoid sharing information with everyone. As these policies scale to hundreds and thousands of rules, however, conflicts will become more difficult to manually detect, which may result in unintended consequences. For example, the right r0 conflicts with the prohibition p0, because the right permits sharing information for the industrial control industry without exceptions, whereas the prohibition prohibits sharing for industrial control with

low impact factors ≤ 4. This kind of conflict is an easy oversight that can be detected using concept subsumption in Description Logic. Authorization languages, such as EPAL [19], XACML [16] and AIR [11, 22] are candidates for evaluation as policy languages.Data Provenance and Quality

Automated decision-making carries privacy risks when data is poor quality or when premises to a conclusion are challenged. Data provenance is a design mechanism to address these risks and consists of maintaining different links, from data to the data supplier and collection authority, and from data to the inferred premises and conclusions about the data. The data supplier link should be sufficient to trace the data back to its origin of storage, such as a webpage or a record in a database. The collection authority must distinguish among several types: if individual consent to monitoring was provided, the date and time that consent was provided and contact information of the consenting subject should be maintained; or if a warrant or subpoena was used, the date and contact information for the issuing authority of the warrant or subpoena should be maintained.

In the event that data is corrected, any conclusions based on inaccurate data must be recalled, if the updated data falsifies premises to those conclusions. This may have cascading effects across multiple systems, which requires system designs to accept and respond to cascading updates. The individual participation principle in the FIPPs allows individuals to correct information about themselves in commercial databases and these databases may be used as input to an automated law enforcement investigation system. Thus, these acts of correction can influence how law enforcement inquiries are predicated, if facts become invalidated over time.

V. Conclusion

Increased access to distributed databases will benefit all-source attribution to identify and prevent emerging cyber attacks. This opportunity increases privacy risk, as this access can include automated decision-making based on non-public and public information. Analogous systems that lack transparency and create tremendous public risk include high-speed, financial trading systems, which are suspected to have caused significant, negative market fluctuations. To manage privacy risk, we discussed important framework elements to consider in regards to inquiry escalation, collection rights and obligations, and data provenance and quality. Areas for further consideration include the role of the Data Mining Reporting Act in developing and reporting escalation patterns and a recent trend by the FBI to respond to cyber security incidents using active defense [24], which is the ability to neutralize an attack or raise the costs to conduct an attack [18].

VI. Acknowledgements

This material is based upon work supported by the U.S. Department of Homeland Security under Grant Award Number 2006-CS-001-000001, under the auspices of the I3P research program.

References

[1] S. Baker, *Skating on thin ice: why we aren't stopping tomorrow's terrorism*. Hoover Institution Press, 2010.

[2] T.D. Breaux, A.I. Antón, "Analyzing Regulatory Rules for Privacy and Security Requirements," *IEEE Trans. Soft. Engr.,* 34(1):5-20, Jan./Feb. 2008

[3] T.D. Breaux, A.I. Antón, J. Doyle, "Semantic parameterization: a process for modeling domain descriptions," ACM Trans. Soft. Engr. Method. 18(2): 5, November 2008

[4] T.D. Breaux, "A Data Flow Control Language for Distributed Systems," *CMU Technical Report ISR-11-103*, March 2011.

[5] F. Baader, D. Calvenese, D. McGuiness (eds.), *The Description Logic Handbook: Theory, Implementation and Applications*, Cambridge University Press, 2003.

[6] C. Dwork, "Differential Privacy," 33rd Int'l Colloquium on Automata, Languages and Programming, part II (ICALP 2006), LNCS v. 4052, 2006, pp. 1-12

[7] B. Fischhoff (chair) et al. *Intelligence Analysis for Tomorrow: Advances from the Behavioral and Social Sciences,* Committee on Behavioral and Social Science Research to Improve Intelligence Analysis for National Security; National Research Council, 2011.

[8] J.F. Horty. "Deontic logic as founded in non-monotonic logic." *Annals of Math. & AI*, 9: 69-91, 1993.

[9] M. Huth, M. Ryan. *Logic in Computer Science, 2 ed.* Cambridge University Press, 2004.

[10] G.V. Jean, "In the Fight Against Cybercrime, Weapons Have Short Shelf Lives," *National Defense*, Sep. 2009.

[11] L. Kagal, C. Hanson, and D. Weitzner, "Using dependency tracking to provide explanations for policy management," *IEEE Policy*, 2008.

[12] O.S. Kerr, "The forth amendment in cyberspace: can encryption create a reasonable expectation of privacy?" *Connecticut Law Review*, v. 33, pp. 503-533, 2001.

[13] C. B. Lotrionte, "Cyber-search and Cyber-seizure: policy considerations of cyber operations and fourth amendment applications," *Law, Policy and Technology: Cyberterrorism, Information Warfare and Internet Immobilization*. Info. Sci. Pub., 2010

[14] C. Lutz. "PSpace Reasoning with the Description Logic ALCF(D)", Logic Journal of the IGPL, 10(5): 535-568, 2002.

[15] A. Machanavajjhala, D. Kifer, J. Gehrke, M. Venkitasubramaniam. "L-diversity: privacy beyond k-anonymity." *ACM Trans. Know. Disc. Data*, 1(1): 3, 2007.

[16] T. Moses, ed. eXtensible Access Control Markup Language (XACML) Version 2.0, Oasis Standard, 1 February 2005. http://docs.oasis-open.org/xacml/2.0/

[17] J. Mueller. *Overblown: How Politicians and the Terrorism Industry Inflate National Security Threats, and Why We Believe Them*. New York: Free Press, 2006.

[18] W.A. Owens, K.W. Dam, H.S. Lin (eds). *Technology, policy, law and ethics regarding U.S. acquistition and use of cyberattack capabilities*, *Nat'l Acad. Press*, pp. 138-141, 2009.

[19] C. Powers, M. Schunter, "Enterprise Policy Authorization Language," Version 1.2, *W3C Member Submission*, Nov. 2003.

[20] L. Sweeney. "k-anonymity: a model for protecting privacy." *Int'l J. Uncertainty, Fuzziness and Knowledge-Based Systems,* 10(5): 557-570, 2002.

[21] S. Warren, L. Brandeis. "The Right to Privacy," Harvard Law Review, v. IV, n. 5, 1890.

[22] K. K. Waterman, S. Wang. "Prototyping fusion center information sharing: implementing policy reasoning over cross-jurisdictional data transactions occurring in a decentralized environment," *IEEE Home. Sec. & Tech. Conf.*, pp. 63-69, 2010.

[23] White House. Cyberspace *Policy Review: Assuring a Trusted and Resilient Information and Communication Infrastructure*. May 26, 2009.

[24] K. Zetter, "With court order, FBI hijacks "Coreflood" botnet, sends kill signal." *Wired Magazine*, Apr. 13, 2011.