Abstract - A government’s ability to quickly collect, analyze, and share information is often considered vital to its national security. At increasing scale, however, this same information sharing is increasingly viewed in tension with the protection of personal liberties. This paper describes a technical method for relieving this tension, offering an approach that enables both to be pursued simultaneously. The approach is founded on the concept of an impenetrable “Black Box” in which information can be placed within, but no person can ever access. Rather, the box is driven by encoded policy statements that define patterns of reasonable suspicion derived from standard of law, as interpreted by a duly constituted policy body. The box is specifically designed to only output detection of authorized patterns, but without ever revealing or enabling any sort of access to the information it contains. With no human involvement in the analytic process, much higher levels of privacy assurance are believed possible.

Keywords: Information Sharing, Privacy

1 Introduction

Two pillars of a democratic society – Security and Liberty – are challenged by the post-9/11 world: How can an open democracy sustain the former without infringing on the latter? A government’s ability to collect, process, analyze, and share volumes of information is commonly regarded as central to its national security. But these needs, driven by a desire to detect and prevent attacks against itself and its allies, increasingly conflict with its constitutional protections of individual liberties.

Current public opinion usually frames this debate as a tradeoff, balancing the sacrifice of some liberties against real or perceived gains in security. No end to this debate is in sight. But this paper posits that security and liberty are not mutually exclusive. Rather, it proposes a method that enables both to be achieved simultaneously, through the careful application of policy and modern technology. This concept is referred to here as Privacy Assurance.

2 Information Sharing

The sharing of information across legal and jurisdictional boundaries enables national security and analytic tradecraft – witness how the 9/11-hijackers were not only connected via airline data and other transactional records, but in at least two cases by threat information already maintained by the U.S. intelligence community. This process has been popularized as “connecting the dots.” But localized information “stovepipes” maintained by individual organizations often are not sufficiently rich in their content to discern the complex network of associations and connections across multiple jurisdictions that realistically describe contemporary threats. In contrast, such patterns often are quickly revealed when these otherwise disparate information sources can be merged and analyzed in aggregate.

Unfortunately, the merging of information sources can quickly exceed the respective authorities of participating organizations, creating new threats to individual liberties and personal privacy.

Alternatively stated, it often may be in the best interests of individual organizations spanning various legal and jurisdictional boundaries to share information, but there may not be adequate trust among the participants, or authority from the citizenry under whom they serve, to allow such sharing. This mistrust can arise from the fear of misuse, insufficient oversight, fear of the exposure of sensitive information, sources, and methods, or the increased risk of unintentional exposure. Trust issues aside, privacy policy in the United States today mandates data minimization – to wit, that civilian agencies should only collect personally identifying information (PII) that is directly relevant and necessary to accomplish the specified purpose of its collection; only retain PII for as long as is necessary to fulfill the specified purpose; and only share data with other agencies when compatible with the purpose for which it was collected. Moreover, U.S. citizens are afforded a fundamental right to be “secure in their persons, houses, papers, and effects, against unreasonable searches.”

The approach advocated here takes the Fourth Amendment to the United States Constitution as a basic
system requirement. Within this framework, U.S. law defines “reasonable suspicion” as the standard of law, based on specific and articular facts and inferences, under which a person may be regarded as being engaged in criminal activities, having been engaged in such activity, or about to be engaged in such activity. Reasonable suspicion is the basis for investigatory stops by the police and requires less evidence than probable cause, the legal requirement for arrests and warrants. Reasonable suspicion is evaluated using the “reasonable person” or “reasonable officer” standard, in which an officer in the same circumstances could reasonably believe a person has been, is, or is about to be engaged in criminal activity.

Such suspicion is not simply a hunch. A combination of particular facts, even if each is individually innocuous, can form the basis of reasonable suspicion. This is pivotal to Constitutional law enforcement and to the method for assuring privacy that is laid out below. It describes how reasonable suspicion can be ascertained from multiple information sources without resorting to unreasonable search. Unreasonable search is interpreted here as any type of investigative process that would reveal information that a reasonable person would regard as private, prior to the establishment of reasonable suspicion or probable cause – and thus protected.

3 Privacy Assurance

So how can reasonable suspicion be ascertained from multiple information sources without resorting to unreasonable search? The approach advocated here posits the existence of a “Black Box.” In this context, a Black Box is a physical (or logical) device whose contents are beyond reach: they can never be examined. The device is specifically engineered so that the information it is fed cannot be revealed to anyone, regardless of authorization, executive privilege, court order, vandalism, or deliberate attack. Information can flow into the Black Box, but once it resides within its boundaries, it can never be accessed. For all practical purposes, the Black Box is impenetrable.

Total impenetrability implies a theoretical extreme that likely would be difficult to achieve, or even more important, to verify or accept in the negative. Consequently, this paper takes impenetrability as the condition in which there exist no known exploitable vulnerabilities that would enable access to the contents of the Black Box. While vulnerabilities may exist, an impenetrable Black Box is one about which a group of reasonable, qualified technical experts will testify that any vulnerabilities inherent in the device’s design have been mitigated, using reasonable techniques to assure its security to within a degree of probability asserted as reasonable by a community of such experts.

What good is a Black Box? Assuming the existence of such a device, it then is possible to “share” information in unique and powerful ways.

Figure 1 below illustrates the basic privacy assurance configuration. At the top center of the diagram is the “Black Box” construct. Across the bottom are representations of independent organizations that span multiple legal and/or jurisdictional boundaries. Each of these organizations, via their respective legal charters, is authorized to maintain a specific body of information, represented by the colored “dot” networks depicted within each. These information “dots” are connected via “links” that represent relationships that the organization has discerned and maintains, consistent with its legal authorization.

Figure 1. The Privacy Assurance “Black Box”

The legal charter of each organization may limit its ability to access or share information and thereby identify corresponding relationships across established boundaries. Sharing this information across such a boundary could in fact constitute a breach of law or, alternatively, a breach of public or congressional trust. Nevertheless, if such organizations were actually able to share their information, new patterns and relationships could be identified from analysis. New patterns of suspicious activity that might impact national security could be identified and acted upon. This information would constitute actionable intelligence.

The solution offered here involves placing relevant information from each contributing organization inside the Black Box. Information can then be connected and processed, but without the possibly of human examination or disclosure. The methods used to do this are well established in contemporary analytic tradecraft: Techniques such as graph analysis can discover relationships among billions of data elements. But if the Black Box is designed to be “non-queryable” by any means, how then can it be of any value?

To address the utility question, the Black Box also has exactly one additional input (on the left in Figure 1) and exactly one and only one output (located on the right). At the left interface, patterns of specific interest are input to the box. These patterns are template-like encodings of generic information relationships that a duly authorized policy body has reviewed and approved for submission into the box. Put another way, the patterns are a set of
analytical rules that define the Black Box’s reasonable search behavior. The only patterns that are admissible to the Black Box are those that the policy body has reviewed and has unanimously confirmed as meeting a certain threshold. In this case, the threshold is the set of observable conditions within the Black Box that meet the legal standard for reasonable suspicion.

Within the Black Box, in addition to the information that it receives from each contributing organization, and the patterns it receives from the policy body, is an algorithm that continuously observes for conditions that match any of the submitted reasonable suspicion patterns. Upon detecting such a pattern, the Black Box outputs an identifier for the pattern and a set of identifiers for the information that triggered the pattern’s detection. This is a continuous process. It is executed in real-time without human intervention, again leveraging current analytic tradecraft. Upon such a detection event, the contributing organizations would be notified of the particular identifiers. They then could investigate further, using their existing analytic capacities and legal structures. If permissible by law, additional information could accompany the output notification to expedite investigation. The specification for such auxiliary information is incorporated into the original pattern definition, enabling the policy body to review and approve, and ensure privacy compliance.

All identifiers output by the Black Box would be available to the policy body or alternatively, to a duly constituted oversight body to continuously verify compliance. In other words, while considerable information is flowing into the Black Box, the only aspect that would ever have external visibility is a reasonable suspicion output. In this manner, organizations and the citizenry they serve can receive the benefits of information sharing, but without exposing information to misuse or the risk of privacy invasion.

The only information that can be submitted to the Black Box is information that at least one or more members of the participating organizations have already been authorized to possess. Similarly, the only information that is ever outputted from the Black Box is that which has been deemed in advance to constitute reasonable suspicion and thus assured to meet the legal standard.

4 Implementation Aspects

The basic design of a Black Box is shown in Figure 2. Everything that flows into and/or out of the box must pass through a carefully designed interface that strictly limits access to a small set of well defined, hardware enforced actions. External data sources at the left side of Figure 2 are connected to the box via a set of input adapters. These adapters transform an input source into a set of unique information items and relationships for representation and processing within the box. These items flow electronically into the box via input converters that transform the input source information into a set of distinct graph elements (i.e. a mathematical node/link structure). These elements are passed across the box boundary via a strict hardware protocol. Within the box, this protocol is processed via a set of interface isolators to ensure that the only operations that can ever be executed are those stated in the box’s formal specification. As part of the implementation process, the design of this converter/isolator set is verified using formal methods to ensure proof-of-correctness of the implementation.

![Figure 2. The “Black Box” Design](image)

At the top of Figure 2, patterns that encode reasonable suspicion policy statements flow into the box via the same adapter-converter-isolator process. Patterns for submission to the box, however, must first be compiled from their source specification language (e.g. an English subset) before being passed through the pattern adapter. Within the box, patterns are again represented as a set of graph elements (i.e. templates).

At the core of the Black Box is the trusted graph-processing engine. This engine is very carefully engineered to ensure that its pattern-matching algorithm cannot be modified in any fashion. Once developed, proven, loaded, and authenticated, this algorithm can never be changed without repeating the entire rigorous, monitored process. This is critical for preventing any type of accidental or adversary-assisted disclosure of private information from the box.

At the right side of Figure 2, patterns that are detected by the processing engine flow out through a similar process, but in reserve order. That is, triggered pattern identifiers and the associated reasonable suspicion information identifiers can only exit the box by passing through an interface isolator. Outside of the box, these items are converted to a form recognizable by an operator or alternatively to a form that can be processed by the contributing source organizations or investigating bodies that participate in the feedback/dissemination loop.

The key aspect of this design is that regardless of what information may flow into the box, the only information that can ever exit is that which was approved and authorized by the policy body as meeting a reasonable suspicion pattern. Furthermore, the box itself is implemented in such a manner that these protections cannot be circumvented via tampering. Hence, the
implementation provides for no back doors, no overrides, no special authorizations, nor any inherent exploitable vulnerabilities, within the limits of the formal verification techniques and assumptions used to specify, design, and engineer its correct operation.

5 Operational Considerations

The privacy assurance approach advocated here is based on existing, well-understood analytic tradecraft and proven, off-the-shelf technology components. While a proof-of-correct-operation is at hand, the full technical aspects of this approach are outside the scope of this paper. However, none of the constituent techniques and components described here are particularly new, distinctly novel, or technically unfounded. It is the careful configuration of these components and their unique operationalization within an existing policy framework that has not been attempted. Consequently, a pilot demonstration program is being pursued to methodically address the intricate blend of policy and technology related issues. These issues include:

- The establishment of a policy body and its associated processes for defining and authorizing patterns that would constitute reasonable suspicion.
- The establishment of an oversight function or oversight body to monitor the operation of a Black Box configuration, including the auditing of input patterns and output notifications to ensure legal compliance.
- The specific configuration, operating procedures, physical and cyber protections, and certifications of the Black Box and its interfaces to ensure sustained operational system integrity.
- The operational polices and procedures for identifying, protecting, and mitigating specific vulnerabilities across an end-to-end system deployment.

6 Conclusion

This paper outlines an approach that enables organizations to share information in a manner that respects individual privacy rights. Although discussed here within a privacy policy context, the Black Box approach is applicable to a diverse spectrum of information sharing challenges:

- As noted, the patterns are simply analytical rules that can be changed like any filter. This means that the Black Box methodology could be applied to facilitate other types of information sharing by intelligence agencies, thus overcoming traditional obstacles to cooperation, such as concerns about protecting underlying data, sources, and methods, as well as varying standards for data and metadata.
- Within a highly compartmented organization, this approach would enable information to be aggregated and analyzed without risk of compartment compromise.
- At an international level, the Black Box approach would enable countries to share and analyze highly sensitive information, perhaps of a treaty or compliance nature, without exposing protected concerns or individual violations.
- Numerous other applications, ranging from health records management and compliance with HIPAA to personnel information processing can be readily envisioned.

As previously mentioned, the strength of this approach rests heavily on the impenetrability assumptions of the Black Box. While specific candidate Black Box implementation techniques exist that offer the highest levels of theoretical mathematical assurance, from a practical standpoint, 100.00% absolute impenetrability assertions are largely unverifiable. Rather, impenetrability must be ascertained via an assessment of risks based on specific threats to components that exhibit varying degrees of trust. Analogous to the legal standard that defines reasonable suspicion, so too would an accompanying technical standard that defines reasonable impenetrability need to be established. A community of concerned scientists, engineers, and technicians from across the country stand ready for this undertaking. The implications of a technical resolution upon the current tension between security and liberty would be profound.

Acknowledgement: This paper is the result of numerous conversations with remarkably insightful individuals from all across the cyber security industry. In particular, significant contributions to this engaging dialog were received from the policy and technology communities at Georgetown University, Lawrence Livermore National Laboratory, and Oak Ridge National Laboratory, and the National Security Agency.