Analytical Tools for Privacy Risks: Assessing Efficacy on Vote Verification Technologies

Rosa R. Heckle
Department of Information Systems
UMBC
1000 Hilltop Circle
Baltimore, MD 21250
Heckler1@umbc.edu

Stephen H. Holden
Department of Information Systems
UMBC
1000 Hilltop Circle
Baltimore, MD 21250
Holden@umbc.edu

ABSTRACT
This study examines the two analytical tools—Privacy Impact Assessments and the classic risk analysis model in an effort to understand their utility and efficacy as tools to assess privacy risk, particularly for assessing the privacy risks of vote verification technologies. Our findings suggest that neither tool alone provides a complete assessment of privacy risks for the technologies examined. The combination of the two, though, effectively helps uncover privacy risks that require mitigation.

1. INTRODUCTION
The Federal government has adopted an analytical framework to assess and mitigate the privacy risks often found in federal information systems. This framework is termed a Privacy Impact Assessment (PIA). While the PIA is prevalent in the public sector, many software engineering firms use an IT risk analysis model to identify threats and vulnerabilities to the systems being designed in an effort to mitigate and manage risks, including privacy [7].

These two models, the PIA and the classic risk assessment, provide a much needed theoretical basis for exploring the privacy risks in software applications. While the PIA was designed specifically to address the privacy risks, the classic risk model is used to assess security risks more broadly. While each of these models offers significant explanatory power, they have a different focus. A model that integrates elements from both may offer a significant improvement over either model alone for assessing privacy risks.

2. MOTIVATION
The State of Maryland was interested in the viability of using vote verification technologies in conjunction with its current Direct Recording Electronics (DRE) solution, the Diebold Accuvote TS. Researchers from UMBC's Maryland Institute for Policy Analysis and Research conducted a technical analysis, including identifying privacy risks, of the vote verification technologies under consideration [6].

Some security experts believe that the core requirements of any valid voting system are security strong enough to prevent fraud and provide the anonymity of a secret vote [5]. However, security is necessary, but far from sufficient, to ensure anonymity of a secret vote. To date, much of the scrutiny of electronic voting systems has focused on security measures on the presumption that if they are secure, then they are verifiable and ballot secrecy is upheld. This could imply that once the security risks have been identified and controlled, privacy is also taken care of.

We believe, however, there is an inherent tradeoff between verification and the threat to privacy; "…with too much information anonymous voting is compromised and too little information prevents effective audit trails" or verifiability [5]. The stricter the verification, the greater the potential disclosure of personal information, and the greater the risk to privacy. For this reason, the security risk analysis model may not be effective in assessing privacy risks, especially for vote verification.

After an exhaustive literature search, I could not find a privacy risk analysis model other than the PIA. Our review of the Human Computer Interaction and privacy literatures suggested the need for one. According to Adams and Sasse, "Most invasions of privacy are not intentional but due to designers' inability to anticipate how this data could be used, by whom, and how this might affect users" [1]. The privacy impact assessment used by the US federal government assists with identifying privacy risks inherent in the system and facilitates mitigation of the risks. Consequently, this model could be effective in identifying privacy risks specifically and could be beneficial in instructing system design. The prevalent model for assessing risks in information systems is the security risk analysis model, whether they are security or privacy risks to be analyzed. In view of this, we have taken both models, the PIA and the classic risk analysis models, and applied them to the vote verification technologies.

3. METHODOLOGY
We completed a privacy risk analysis of four vote verification technology products being evaluated by the State of Maryland. Additionally, the analysis identified which of the two models identified the privacy risks. Outlined below are the details of the models, the systems reviewed, and the procedures.

Privacy advocates have adopted the general idea of an impact assessment, defined as "the identification of future consequences of a current or proposed action" [2] and have modeled it particularly for addressing privacy issues. The purpose of conducting a privacy impact assessment is to analyze how personal information is collected, used, stored, and protected in information systems. Recent federal legislation, the E-government Act of 2002 created a requirement for federal agencies to conduct PIAs for selected information systems.

Beyond PIAs, both public and private sector organizations have used IT risk analysis models to identify threats and vulnerabilities

to the systems being designed in an effort to mitigate and manage risks. A risk analysis is the process of examining a system for possible problems that can arise and then planning to mitigate the effects of those problems [7]. The classic risk assessment model used for this analysis was taken from the National Institute of Science and Technology [8].

This study included a review of four separate verification technologies. These included the following organizations and individuals: VoteHere (Sentinel); SCYTL (Pnyx.dre); Prof. Ted Selker, MIT (VVAATT); and Diebold's VVPAT.

The first step was to identify each verification system's characteristics and operating environment. In this analysis, the vote verification systems were not viewed in isolation but within the context of the entire voting process and the conduct of elections. The scope of the review looked specifically at the verification module as it would be integrated with the Diebold DRE within the policies, procedures, and laws of the State of Maryland. The data used in this assessment came from a variety of sources, including data and documentary material provided by the vendors, and the State of Maryland Elections policies and guidelines specified in the Code of Maryland Regulations.

The second step was to take each system through the PIA process, which means completing a PIA document. This research used the PIA document used by the Internal Revenue Service[2], which is a tool that facilitates an in-depth look at the data being collected, used, accessed, and shared and ultimately protected. We addressed all the questions posed in the PIA document and then reviewed the findings to identify and highlight the privacy risks.

The third step was to evaluate each system using the classic risk analysis model. The risk assessment methodology defined by NIST encompasses nine primary steps: System Characterization, Threat Identification, Vulnerability Identification, Control Analysis, Likelihood Determination, Impact analysis, Risk Determination, Control Recommendations, and Results Documentation. We took the assessment through to the risk determination step. To address the steps of Vulnerability Identification, Control Analysis, Likelihood Determination, Impact analysis, we informally considered a number of metrics taken from the NIST assessment model and from the Preliminary Threat Analysis to voting machines. The metrics included: risk tolerance, size and diversity of the conspiracy, resources needed, and impact. Three measures were used; low, medium, high. High Risk tolerance is taken to mean the attack has a higher risk of being exposed, and low risk tolerance suggests that the risk could go without being noticed [8].

4. FINDINGS AND DISCUSSION

The privacy risk rating for each of the vote verification technologies analyzed was based on two factors: the control mechanisms in place and the sensitivity of the data that was being stored. Neither the PIA nor the classic risk assessment model was effective in reviewing both factors and disclosing the full spectrum of privacy risks. The results suggest that any privacy risk assessment model must review both factors to be effective.

The classic risk assessment model, as applied, did not address the privacy sensitivity of the data, nor the privacy considerations in the collection, storage, and use of the data within the verification systems. Though a key component of the risk assessment is the identification and classification of assets (where data are considered an asset) it does not provide guidelines on how to classify data in accordance to its privacy sensitivity. This is where the PIA could be very useful. The questions in the PIA document specifically address data, who would have access to it, storage and retention periods, etc. in detail.

The Privacy Impact Assessment, on the other hand, did a good job of identifying the data sensitivities of the verification systems, but did not identify the full spectrum of threats to that system. Again, the PIA did ask, "What controls are in place to prevent the misuse (e.g., browsing) of data by those having access? What controls will be used to prevent unauthorized monitoring" [4]. These questions could be construed as asking what security mechanisms are in place to prevent unauthorized access. However, it does not address all threats or guide the evaluator to assess other threats to the system.

Each method assesses risk from different perspectives. The core of the PIA is a careful description of how a system actually works like the standard operating procedures or process. The classic model, on the other hand, has at its core the assessment of threats to the system. While the PIA determines if there are any privacy risks inherent in the system design, the classic model determines what can go wrong with the system that would threaten privacy.

5. FUTURE RESEARCH

For assessing risk in the voting verification systems, the need to blend the classic risk assessment model and the Privacy Impact Assessment model was necessary to identify and deal effectively with the full spectrum of threats to privacy. Further work needs to be done to develop and finalize a holistic privacy risk analysis model, and then empirically test it on different applications to determine if the new model can be instrumental in assessing privacy risks inherent in new system deployments.

6. REFERENCES

[1] Adams, A., and Sasse, A., Privacy in multimedia communications: protecting users, not just data, online: http://www.cs.ucl.ac.uk/staff/A.Sasse/adamshci2001.pdf

[2] Internal Revenue Service Model Information Technology Privacy Impact Assessment, online: www.cio.gov

[3] Clarke, Roger, Privacy Impact Assessments, 2003, online: http://www.anu.edu.au/people/Roger.Clarke/DV/PIA.html

[4] Internal Revenue Service Model Information Technology Privacy Impact Assessment, online: www.cio.gov

[5] Keller, A., et al., (2004), Privacy Issues In An Electronic Voting Machine", in ACM Workshop on Privacy in the Electronic Society, pages 33-34, Online: http://www.sims.berkeley.edu/"jhall/papers/

[6] Norris, Donald, F., et al, A Study of Vote Verification Technologies, Part 1, Technical Study, Prepared for the Maryland State Board of Elections, 2006

[7] Pfleeger, Charles P., Security in Computing, Prentice Hall, Third edition, 2003

[8] Stonebumer, G, et al, A, Risk Management Guide for Information Technology Systems, Recommendations of the National Institute of Standards and Technology, online: http://csrc.nist.gov/publications/nistpubs/800-30/sp800