Comments on NIST Privacy Engineering Objectives and Risk Model

Charlie Williamson

Legal Counsel, Intelligence Systems Support Office

SAF/ISSO, 301-203-2671

charles.w.williamson12.civ@mail.mil

Summary

- NIST is proposing a privacy engineering model: http://csrc.nist.gov/projects/privacy_engineering/privacy_engineering_presentation_sept_2014.pdf
  - More at http://csrc.nist.gov/projects/privacy_engineering/index.html
- The NIST model provides valuable factors in assessing privacy interests, but does not account for legitimate governmental interests on behalf of society, like national security, national defense, and law enforcement
- The model needs a broader approach to be useful as a decision tool for assessing privacy measures
- Based on its wording, some relying on the current model could conclude that "less data is better" and "less functionality is better" in all cases
- These notes propose a broader model following a standard benefit-cost approach adapted to capture NIST privacy model features
- These notes also identify specific concerns in NIST's Privacy Engineering Objectives and Risk Model

Disclaimer

- These notes represent the thoughts of the author and are not the coordinated or official position of the US Air Force or any other Department of Defense entity.

NIST Model

Key Features and Concerns About Them

- Key features include:
  - Risk management framework and engineering components
  - Privacy objectives: the “privacy triad”
  - System Privacy Risk Equation
- The key concerns are that:
  - The risk model is not parallel to typical risk approaches, so it can be confusing
  - Factors of the risk model cannot be evaluated, so options cannot be compared

Model Privacy Risk Management Framework

Policy
(Law, Regulation, FIPPs, etc.)

Risk Assessment

Requirements

System

Evaluation

Privacy Engineering Components

Objectives
(Predictability, Manageability, Confidentiality)

Risk Model
(Personal Information + Data Actions + Context = System Privacy Risk)

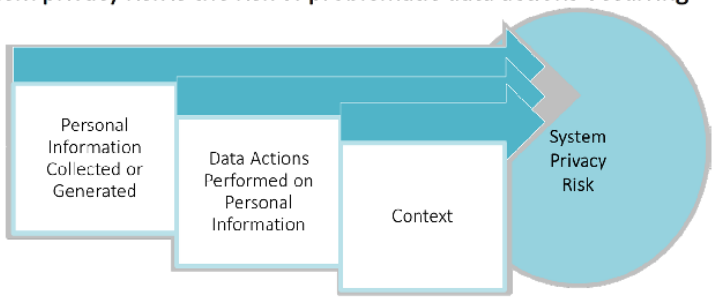
Controls
(Derived from FIPPs, etc.)

Metrics

September Workshop

System Privacy Risk Equation

System privacy risk is the risk of problematic data actions occurring



Personal Information Collected or Generated * Data Actions Performed on that Information * Context = System Privacy Risk

NIST

The Privacy Triad

- The objectives are characteristics of the system, not role-based.
- The objectives support policy
- Aligning the privacy and security overlap

Predictability
Enabling reliable assumptions about the rationale for the collection of personal information and the data actions to be taken with that personal information.

Manageability
Providing the capability for authorized modification of personal information, including alteration, deletion, or selective disclosure of personal information.

Confidentiality
Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. (NIST SP 800-53, rev 4)

NIST

10

Risk Models Generally

Risk Assessment and Risk Mitigation

- Risk assessment is plainly different from risk mitigation
- Risk assessment:
  - Risk is often calculated as the magnitude of loss from an event times its probability of occurrence
  - See, for instance, EPA Risk Assessment: "While there are many definitions …, EPA considers risk to be the *chance* of *harmful effects* to human health or to ecological systems …" http://epa.gov/riskassessment/basicinformation.htm#risk
- Risk mitigation:
  - By comparison, decisions about risk are then often based on Benefit-Cost Analysis, in which decisions are made by comparing the value of a system before a proposed change to the value of a system expected after the change, in light of the cost
- A broader approach is described in ISO 31000 'Risk management – Principles and guidelines,' but it still incorporates the elements above. See, e.g., p. 5, http://www.theirm.org/media/886062/ISO3100_doc.pdf

Calculating Risk and Mitigation

- As a calculation, where R_T = total risk, L_i = loss from single event, and $p(L_i)$ = probability of loss from that event, then

$$R_T = \sum_{i=1}^{n} L_i * p(L_i)$$

- Where State 1 (S_1) describes a system before mitigation, State 2 (S_2) describes a system after mitigation, and C is the total cost of each respective state, then

given $S_1 = \frac{R_{T1}}{C_1}$ and $S_2 = \frac{R_{T2}}{C_2}$,

mitigation is appropriate if $S_2 > S_1$.

Risk Acceptance

- In society, some people are allowed to accept risk for themselves, while some risks to innocents or bystanders must be mitigated
  - Adults making self-determined, fully-informed choices about risk to themselves generally proceed with minimal external regulation even for inherently dangerous activities (i.e., sport parachuting, semi-pro rodeo), but sponsors may still do risk evaluation and mitigation
  - Risks imposed by designers on users without full choice, or on users external to the system, are subject to legal regulation or legal consequence, i.e., building codes or motor vehicle safety are regulated by govt, by liability, and the insurance market
  - Risks imposed by designers on innocent users are subject to enhanced legal regulation or legal consequence, i.e., child safety seats get federal regulation; smoking is regulated for minors

Inside and Outside Risks

- Risks to people inside and outside the system should be considered
  - "Authorize information system operation *based upon a determination of the risk to organizational operations and assets, individuals, other organizations and the Nation* resulting from the operation of the information system and the decision that this risk is acceptable." NIST Risk Management Framework (RMF) OVERVIEW, http://csrc.nist.gov/groups/SMA/fisma/framework.html
  - "Major Non-Conformity (MNC) … poses a *serious threat to* the safety of *personnel* or the *ship or* a serious risk to the *environment …"* The Maritime Safety Committee of the International Maritime Organization … DEFINITIONS (1.1.10), http://www1.veristar.com/veristar/dps_info.nsf/veristar/Dps_info.nsf/AllByDateInternal/353FC02A22D59873C1257734003EC51C?opendocument
  - EPA Risk Assessment: "While there are many definitions of the word risk, EPA considers risk to be the chance of harmful *effects* to *human health or* to *ecological systems …"*
- Focusing only on "system risk" leads to under-appreciation of total risk
  - For instance, power plant and auto engineers did not consider global warming until recently because they treated air and water as unlimited sinks

Risks and Costs

- All risks should be considered in light of utility and cost
  - "Incorporating the same concepts used in managing information security risk, helps organizations implement privacy controls in a more *cost-effective*, risked-based manner …" NIST Special Publication 800-53 Revision 4, Security and Privacy Controls for Federal Information Systems and Organizations, http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r4.pdf
  - See also, NIST Special Publication 1082, A Guide to Printed and Electronic Resources for Developing a *Cost-Effective* Risk Mitigation Plan for New and Existing Constructed Facilities
  - "… overall risk management should be a fundamental driver of an organization's approach to privacy: solutions should be risk-based and *affordable*." http://www.nist.gov/cyberframework/upload/privacy-workshop-summary-052114.pdf
  - Consider evolution of motor vehicle safety over time: society has increasingly valued safety, but not so much that everyone drives tanks at 3 mph
- Utility and cost are assessed both for the individual and for society
- Some privacy is positive for some and negative for others, i.e., grand jury hearing secrecy is positive for indictees and negative for media

Benefit-Cost Analysis

Benefit-Cost Analysis is Widely Used and Appropriate for Privacy Analysis

- Many models are available
- It is widely used across the govt
  - OMB: http://www.whitehouse.gov/omb/circulars_a094#5
  - FEMA: http://www.fema.gov/benefit-cost-analysis
  - US Army: http://asafm.army.mil/Documents/OfficeDocuments/CostEconomics/guidances/cba-gd.pdf
  - Dept of Justice: https://ncjrs.gov/pdffiles1/nij/246769.pdf; http://cbkb.org/about/
- It can be used to inform corporate risk management and insurance decisions, so privacy approaches using this Analysis are more likely to be adopted than approaches that do not consider cost and utility
- It is cross-disciplinary: it is used and understood by engineers, accountants, executives, and regulators
- It can be applied across organizations: use of similar models allows entities to show how flows of data between them affect overall privacy
- Because it is used at the point of investment decisions, it encourages early consideration of security-by-design and privacy-by-design principles
- It can even be used to build competitive advantage by showing rigor in privacy protection
- It is not perfect; some cautions are appropriate: see, e.g., http://www.econlib.org/library/Enc/BenefitCostAnalysis.html

Building on the NIST Model

- The NIST model proposes:
  - System Privacy Risk = Personal Information Collected or Generated * Data Actions Performed on that Information * Context
  - The discussion deck also gives a number of factors that relate to each component
- The disadvantages of this approach are that:
  - There is no consideration of utility of the system, so privacy is treated as the only value
  - In the absence of other guidance, there will be a natural tendency to minimize risk, which means risk is simply best managed by pushing each component toward zero
  - In plain language, "less data is better" and "less functionality is better"
  - This is not how info system design works
  - There is no consideration of cost, so commercial entities will have to figure out how to factor that into their investment decisions, and govt entities are required to do so for regulatory decisions, Regulatory-Right-to-Know Act, Section 624 of the Treasury and General Government Appropriations Act of 2001, Pub. L. No. 106-554, 31 U.S.C. § 1105 note.
  - It is contrary to most engineering approaches which try to optimize systems rather than minimize a component

Model Application

- Whether an approach like NIST's or a benefit-cost approach is used, the model chosen needs to be able to guide decisions on everyday scenarios like financial transactions, on-line purchases, mobile applications, and the Internet of Things
- It also needs to aid decisions on outlier scenarios like these:
  - What if the chance of curing major cancers was proportional to access to the following info: individual gene maps, addresses for the last 15 years, and income history? What is the effect of individual opt-out?
  - What if the chance of finding a child kidnapper relied on quick access to a toy store database?
  - These kinds of scenarios may require societal level debate, but consensus on a model will aid discussion

Benefit-Cost Proposal: Narrative

- System changes should be made when benefits exceed costs
- Info System Benefits = Optimized (Utility + Security + Privacy)
- System Costs = Creation + Compliance + Enforcement
  - Creation Costs = Costs to design and implement utility, security, or privacy measures
  - Compliance Costs = Costs to comply with or maintain utility, security, or privacy measures
  - Enforcement Costs = Costs to oversee, inspect, regulate, and litigate implementations of utility, security, or privacy measures

Features of a Benefit-Cost Approach

- Combining Utility, Security, and Privacy considerations drives analysis of trade-offs
  - For instance, more utility may mean less privacy, and analysis could show the change should not be made
  - Or, a stronger privacy feature may allow less investment in security
- The goal is to optimize rather than balance: what is the best combination?
- Utility and Security benefit/cost models exist and need not be created
  - Utility benefit/cost is generally based on projected sales for commercial work or return on investment for govt work. See, e.g., http://portal.hud.gov/hudportal/documents/huddoc?id=DOC_15127.doc
  - Existing sample security benefit/cost models: http://www.sei.cmu.edu/reports/04tn045.pdf, http://www.courant.nyu.edu/ComplexSystems/literature/Arora,etal.pdf
- However, advances are needed in Privacy benefit-cost models
- Benefits and costs should be summed for all persons affected by the system
  - An increase in a criminal's privacy could lead to an decrease in utility for past victims, potential victims, and law enforcement
  - An increase in an activist's privacy could be a positive for democracy
- Decisions will not be apolitical or globally applicable in all markets or regions: an activist in one society will be an agent provocateur in another
- It will take some experience and data to quantify benefits/costs and optimize

Privacy Benefit and Cost Factors

Privacy Benefits Could Include:	Privacy Costs Could Include:

• Predictability : known or reasonably expected disclosures of personal information. Protects against:
  • Induced Disclosure
  • Unanticipated Revelation
  • Unauthorized Surveillance
• Manageability: disclosures of correct personal information for legitimate personal or societal purposes. Protects against:
  • Distortion
  • Discrimination
  • Unwarranted Restriction
(See http://csrc.nist.gov/projects/privacy_engineering/privacy_engineering_presentation_sept_2014.pdf for descriptions of terms)
	• Creation of a new privacy control
	• Compliance with privacy protections by system owner
	• Enforcement of privacy protections by the system owner, govt, or other external authority (i.e., insurers)

Additional factors may be appropriate.

Benefit-Cost Proposal: Evaluative

- Where:
  - State 1 (S_1) describes a system before change and State 2 (S_2) describes a system after change
  - B_{T1} = Total benefits of State 1; B_{T2} = Total benefits of State 2
  - C_{T1} = Total costs of State 1; C_{T2} = Total costs of State 2

  Then, given $S_1 = \frac{B_{T1}}{C_{T1}}$ and $S_2 = \frac{B_{T2}}{C_{T2}}$, change is appropriate if $S_2 > S_1$.
- Adding the NIST privacy factors yields

  $$S_\alpha = \frac{B_{T\alpha}}{C_{T\alpha}} = \frac{\sum(B_{Utility\alpha}, B_{Security\alpha}, B_{Privacy\alpha})}{\sum(C_{Utility\alpha}, C_{Security\alpha}, C_{Privacy\alpha})}$$, where

  $$B_{Privacy\alpha} = \sum_{i=1}^{n} Predictability_i + \sum_{i=1}^{n} Manageability_i$$

  $$C_{Privacy\alpha} = Creation + \sum_{i=1}^{n} Compliance_i + \sum_{i=1}^{n} Enforcement_i$$

  i= people affected by the system, internally or externally
- Some factors may be very difficult to evaluate, so they may have to be assessed instead, but the method still provides a disciplined, repeatable approach that can develop reusable modules for analogous systems over time.
- Next step: derive algorithm to optimize S_α rather than just evaluating it

Use Case – Privacy Benefit/Cost
Two-person authentication for Sysads

- Assume good security practices (i.e. least privilege, intrusion prevention, good passwords, etc.)
- What is the analysis for preventing disclosure of a particular user's true name as being associated with a commercial social media application?
- Proposed privacy measure: add two-person authentication for sysads to release true names
- Assume 1M users; 1K with "sensitive" names (i.e., celebrities, political figures, activists)
- Assume sensitive users would be willing to pay an extra $500 for this feature but others would only be willing to pay $.10
- Assume it costs $5K to create, $5K/yr to internally enforce and a 10 yr life, $0 for external enforcement

Use Case – Privacy Benefit/Cost

Two-person authentication for Sysads

$S_\alpha = \frac{B_{T\alpha}}{C_{T\alpha}} = \frac{\sum(B_{Utility\alpha}, B_{Security\alpha}, B_{Privacy\alpha})}{\sum(C_{Utility\alpha}, C_{Security\alpha}, C_{Privacy\alpha})}$, where

$B_{Privacy\alpha} = \sum_{i=1}^{n} Predictability_i + \sum_{i=1}^{n} Manageability_i$

$C_{Privacy\alpha} = Creation + \sum_{i=1}^{n} Compliance_i + \sum_{i=1}^{n} Enforcement_i$

$B_{Privacy2} =$
$\sum_{i=1}^{n}((\$500/\text{person} * 1{,}000 \text{ sensitive users}) + (\$.10/\text{person} * 999{,}000 \text{ other users})) + \sum_{i=1}^{n}(\$0/\text{user} * 1\text{M users})$
$= \$599{,}900$

$C_{Privacy2} = \$5{,}000 + \$50{,}000 + \$0 = \$55{,}000$

$S_2 = \frac{\$599{,}000}{\$55{,}000} = 10.9$, so the social media company ought to adopt this proposed privacy measure

- The benefit-cost ratio must only be >1 for the measure to be appropriate, and need not be compared to a S_1 because there are no utility or security tradeoffs to consider

Use Case Notes

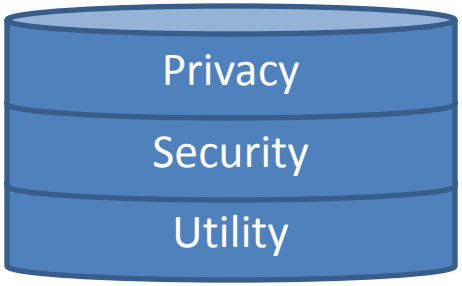
- This example also hints at how simple it would be to automate this process and to do sensitivity analyses of assumptions and entries with simple substitutions, i.e.,
  - What would be the analysis for getting user consent for release?
  - What would be the analysis if such a practice was regulated?

Specific Feedback on the Current NIST Model

- The current NIST model has a number of drawbacks
- The following slides discuss some of them

Slide 6 - Scope

- CURRENT
  - Security and Privacy are shown as intersecting circles
- RECOMMEND
  - Show them as separate zones of interest (e.g., pillars or columns), but have security support privacy, and add utility:



- RATIONALE
  - Info systems need all three to be successful, and intersecting security and privacy may cause some to consider factors like confidentiality twice, when it is already fully considered under the Confidentiality-Integrity-Availability security model

Slide 10 - The Privacy Triad

- CURRENT
  - The slides propose a triad of Predictability, Manageability, and Confidentiality
- RECOMMEND
  - Consider Confidentiality under security
  - Define Predictability more broadly: Enabling reliable assumptions about the rationale for the collection of, access to, and disclosure of, personal information and the data actions to be taken with that personal information.
- RATIONALE
  - This better aligns with Fair Information Practice Principles and eliminates double consideration of Confidentiality

Slide 13 - System Privacy Risk Equation
Slide 14 - Context

- CURRENT
  - System Privacy Risk = Personal Information Collected or Generated * Data Actions Performed on that Information * Context
  - "Context" means the circumstances surrounding a system's collection, generation, processing, disclosure and retention of personal information.
- RECOMMEND
  - Adopt a benefit-cost approach integrated with utility and security

- RATIONALE
  - To minimize risk, users need only eliminate data (drive collection or generation to 0), or eliminate functionality (drive data actions to 0), which would also drive the need for the system to 0
  - Also, it is unclear how "Context" would be evaluated under the current defn. Would more "circumstances surrounding" make privacy better?
  - The privacy model seems to rely on the Security Risk model on Slide 12, but that contains both loss (Impact) and probability (Vulnerability and Threat) factors, while the privacy model does not
  - The current model also has no way to account for the value of various types of personal data. My name is unremarkable, my health history should be protected, only my wife and my bank should see my finances absent a court order, and my voting record is sacrosanct even from a court. The current model just treats data as data.

Slide 16

- CURRENT: Privacy engineering is a collection of methods to support the mitigation of risks to individuals of loss of self-determination, loss of trust, discrimination and economic loss by providing predictability, manageability, and confidentiality of personal information within information systems.
- RECOMMEND: Privacy engineering is a collection of methods to support the management of risks to provide predictability and manageability of personal information within information systems.
- RATIONALE:
  - As mentioned before, not all risks are to be mitigated, but all risks are to be considered and managed.
  - The focus on management rather than mitigation is also more consistent with ISO 31000.

Next Steps

- Request that NIST consider another workshop on the basis of publicly-submitted comments