L. Sweeney. Privacy-Preserving Surveillance using Selective Revelation. Carnegie Mellon University, LIDAP Working Paper 15, February 2005. Updated journal version, to appear October 2005. (Updated version available as IEEE Intelligent Systems Sept-Oct 2005.)

Privacy-Preserving Surveillance using Selective Revelation

Latanya Sweeney

Data Privacy Laboratory, School of Computer Science
Carnegie Mellon University, Pittsburgh, PA 15213-3890
latanya@cs.cmu.edu

Abstract

Following the events of September 11, 2001, many in the American public falsely believe they must choose between safety and privacy. This paper proposes an approach to technology (termed "Selective Revelation") that allows data to be shared for surveillance purposes such that shared data have provable assurances of privacy protection while remaining practically useful. Data are provided to a surveillance system with a sliding scale of identifiability, where the level of anonymity matches scientific and evidentiary need. During normal operation, surveillance is conducted on sufficiently anonymous data that is provably useful. When sufficient and necessary scientific evidence merits, the system drills down increasingly more identifiable data. This is a computational model of the "probable cause predicate" performed in American jurisprudence. Under Selective Revelation, human judges, who make decisions as to whether information will be shared with law-enforcement, are replaced with technology that makes these decisions for broader surveillance purposes.

Introduction and Problem Statement

Society is experiencing exponential growth in the number and variety of information collected on individuals [1]. As the price of disk storage continues to plummet, the cost of capturing and sharing data approaches zero, making it economical to increasingly capture more and more information on the daily lives of individuals. It is not surprising that organizations collect more person-specific information than ever before, and often do so without any particular purpose [1].

When fragments of captured information are combined, they provide person-specific, population-based data ("Databases") that profile individuals. It is believed Databases may reveal behavioral patterns of individuals engaged in illegal activity or forthcoming terrorism. Therefore, one proposed use for Databases is homeland security, which combines law-enforcement and intelligence surveillance activities (together termed "Surveillance").

Examples of American programs that sought to use Databases for Surveillance include CAPS II and TIA [2], both of which faced serious turmoil due to privacy concerns.

Five Privacy Concerns

Databases used for Surveillance summon a myriad of privacy concerns, including the following five.

1. Most people whose information is in the Database have done nothing wrong to warrant suspicion.

2. Surveillance on Databases tends to exasperate privacy expectations and personal protections. While American courts have historically ruled that a person in a public space should have no expectation of privacy [3], information stored in Databases can be so invasive as to remove private enclaves typically available within public spaces. For example, on a crowded bus, one can orient a document to limit what others may see. But, it is difficult to limit views of the document from a hidden camera having a zoom lens because its existence and viewing angle are unknown.

3. Information in Databases can be gathered from private spaces. For example, a private inquiry made on a home phone can become part of a Database, making it indistinguishable from inquiries made at a public counter.

4. Organizations using Databases for surveillance purposes do not implement Fair Information Practices [4] because of a belief that criminals and terrorists may alter their information or behavior. Therefore, no individual whose information is included in a Database has control over his information. No consent is sought. No notice is given. Typically, the subjects of the information do not know their information is being held, and there is no right to or means of correction in cases where information may be incorrect.

5. There is no judicial review or impartial oversight to weigh societal benefits against individual risks. No independent third party limits fishing expeditions or unwarranted inquiries.

The goal of the work presented herein is to guarantee (or at least maximize) privacy protection while making data useful for Surveillance. This work introduces a framework for using Databases such that: (1) no person whose information is contained in the Database can be re-identified without permission; (2) investigators can access necessary information contained in the Database freely and easily; and, (3) results from qualified inquiries are equivalent to results found in the absence of privacy protection. These are termed the "Privacy Conditions for Databases."

Privacy Conditions for Databases

1. No person whose information is contained in a Database can be re-identified without permission.
2. Investigators can access necessary information contained in a Database freely and easily.
3. Results from qualified inquiries are equivalent to results in the absence of privacy protection.

Figure 1. Privacy conditions for person-specific databases used for surveillance.

Methods

One way to satisfy the Privacy Conditions for Databases (see Figure 1) is to technologically model the probable cause predicate in American jurisprudence. When a law officer wants to intrude on a person's private life or affairs, she needs a search warrant, which may be issued by a human judge. In the general case, an officer appears before the judge and reports either facts for which she has first-hand knowledge or facts that she was told through an informant. Typically, the judge in making a decision uses a two-prong test to answer: (1) what is the basis of the knowledge; and (2) is the source believable. See Figure 2 (top). This process can be modeled in technology by replacing the officer with anomaly or data mining algorithms, and the informant with data provided from various data sources. The human judge is replaced with a combination of contracts with the original data collectors and a technologically-enforceable policy statement having preset levels to match the identifiability of the provided information with the minimal information needed by the algorithm. See Figure 2 (bottom). The technology capable of enforcing the policy is called "Selective Revelation."

The first step to construct a Selective Revelation System requires identifying the algorithms to be used and the kinds of data involved. Analyses must then be performed to provably anonymize the data[1] and to then verify that the algorithms remain useful with the anonymized data.

Once the initial step is completed, related regulations, policies, best practices, and laws are mapped onto the scale of identifiability, from anonymous to identified, to specify the authority by which data may be accessed at each status. See Figure 4 (right). Finally, boundaries of algorithmic utility are established to identify the algorithmic circumstances under which more identifiable data is needed. These steps are summarized in Figure 3.

[1] One way to provably anonymize data is k-anonymity [6], but k-anonymity is not the only way. Another example is Privaert [7]. Many other ways are possible.

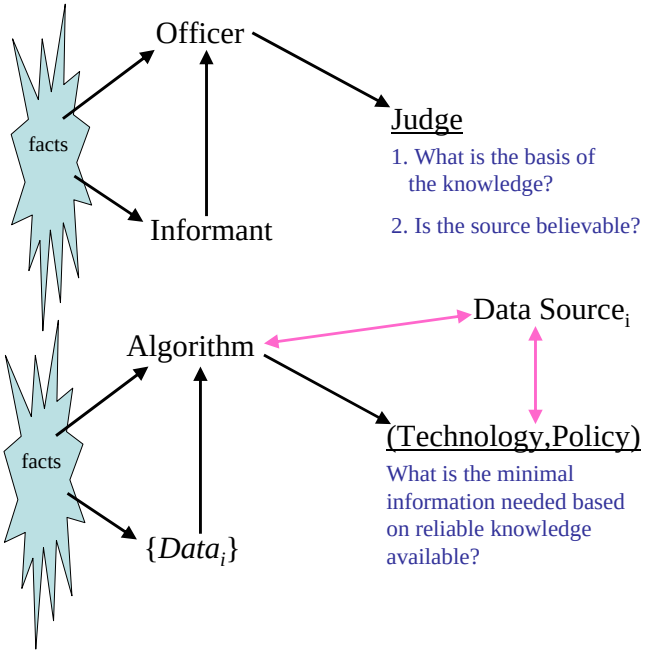

Figure 2. Probable cause predicate as conducted by a human judge (top) and by technology (bottom).

Constructing a Selective Revelation System

1. List algorithms (or class of algorithms) known to be useful.
2. Provably anonymize data.
3. Prove algorithms remain useful with anonymized data.
4. Map operational access constraints specified by regulations, laws, and practices onto identifiability scale.
5. Establish utility boundaries for algorithms with respect to identifiability of data and map onto scale of investigation status.

Figure 3. Steps for building a Selective Revelation System to control Database access.

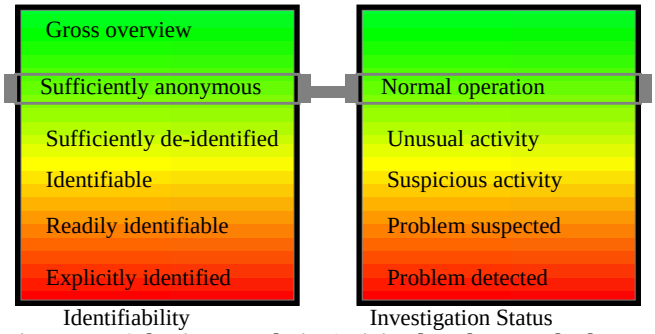

Figure 4. Selective Revelation's joined scales match the identifiability of the data (left) to the operational status of the algorithm used in the investigation (right). Under normal operation, sufficiently anonymous data is used. As suspicious behavior is detected, the investigation status lowers, releasing more identifiable data.

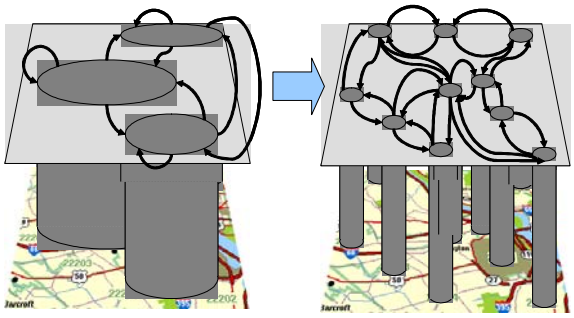

Figure 5. Dynamically augment data access as surveillance warrants. Crude relationships derived from "sufficiently anonymous" data (left). More details revealed using "identifiable" data (right).

Figure 4 shows how identifiability maps to investigation status. During normal operation, anonymized data is used. If unusual activity is encountered, then the identifiability of related cases is lowered to "de-identified" (which has no explicit identifiers, but is not provably anonymous). As the investigation status shifts downward, the provided information becomes increasing more identifiable, until the criteria is met for providing explicitly identified data. Figure 5 demonstrates the effect of lowering identifiability. As the anonymity is lowered, from the right image to the left, more detailed information is made available.

During the operation of Selective Revelation, immutable (tamper resistant) logs are maintained enabling hindsight review of what information was requested, who requested it, the purpose intended, and the version of the data provided.

Application Example

A Selective Revelation System for bio-terrorism surveillance was constructed in which medical data was provided by hospitals, physicians and labs to a public health agency to determine whether an unusual number of respiratory cases were presented [5]. The data were rendered anonymized under the scientific standard of the medical regulation known as HIPAA using Privacert Compliance [7]. The Early Aberration Reporting System algorithm from the Centers for Disease and Control worked with the anonymized data as well as it has done in the absence of privacy protection. If evidence of unusual activity was found, anonymity was lowered to provide more information as to further determine whether the anomaly was most likely an outbreak. If evidence emerged that an outbreak was underway, fully identified data was provided under Public Health Law. This example serves to demonstrate how the American public can enjoy both safety and privacy using Selective Revelation.

Discussion

While Selective Revelation is modeled after the probable cause predicate, it can be used in many cases where search warrant protection is not the governing rule. In the described application for bio-terrorism surveillance, for example, medical data protected under HIPAA is provided. This does not imply that access to medical data must adhere to search warrant protection. Selective Revelation is not limited to applications requiring search warrants. Instead, Selective Revelation models the process of search warrant issuance in order to provide limited data access, and is particularly useful in Database settings not currently restricted by regulations or laws.

Selective Revelation differs from the probable cause predicate in that decisions are based on a sliding scale of identifiability and not a binary one. When a human judge makes a search warrant decision, the result is typically binary –access is granted or not. But under Selective Revelation, the result is nuanced. The decision determines which version of the data will be provided (from anonymous to explicitly identifiable), not whether data will be provided at all.

In terms of the five privacy concerns that motivated the Privacy Conditions for Databases, Selective Revelation provides impartial, automated oversight to all Database inquiries. It does nothing to thwart the collection processes that lead to the existences of Databases. It can be used to help implement some Fair Information Practices by providing hindsight information from its logs that reveals what information was provided on which individuals. But overall, the strength of Selective Revelation stems from its ability to limit data access and add accountability while enabling data use.

References

[1] Sweeney, L. Information Explosion. *Confidentiality, Disclosure, and Data Access*, Zayatz, et al. (eds), Washington, DC, 2001.

[2] *Big Brother is watching you*. World Net Daily, January 16, 2003. www.worldnetdaily.com/news/article.asp?ARTICLE_ID=30523.

[3] Rest. 2d, §652B.

[4] *Fair Information Practice Principles*. U.S. Federal Trade Commission. www3.ftc.gov/reports/privacy3/fairinfo.htm

[5] Sweeney, L. Privacy-Preserving Bio-terrorism Surveillance. *AAAI Spring Symposium*, AI Technologies for Homeland Security, 2005.

[6] Sweeney, L. k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10 (5), 2002; 557-570.

[7] Privacert, Inc., www.privacert.com.