Privacy-Enhanced Linking

Latanya Sweeney
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA USA
latanya@privacy.cs.cmu.edu

ABSTRACT

While computer scientists are uniquely situated to incorporate privacy protections in the link analysis algorithms they construct, most computer scientists are unaware of this opportunity and of ways to think about achieving needed protections. The work presented in this writing introduces a new way for computer scientists to think about providing privacy protection within link analysis and introduces the notion of "privacy-enhanced linking" as algorithms that perform link analysis with guarantees of privacy protection modeled after the Fair Information Practices. In this approach, privacy protection is realized by assessing the validity and interpretation of link analysis results such that inappropriate harm to individuals is provably minimized.

Keywords

Privacy-enhanced technology, privacy, link analysis

1. INTRODUCTION

While law enforcement and counter-terrorism objectives[1] encourage the development of algorithms that learn sensitive information from volumes of disparate data left behind as people conduct their daily affairs, the potential for serious harm to innocent individuals evokes grave privacy concerns.

More specifically, society has experienced exponential growth in the number and variety of data collected on individuals [6]. This growth has been driven by access to inexpensive computing devices and by the plummeting costs of data storage. The ability to collect more data has impacted policy. Throughout the 1990's a pattern emerged in American policy in which policymakers responded to many pressing issues by expanding existing data collections and by starting new data collections on the populace[6].

Having so much readily available data on so many individuals has since ignited a new behavioral pattern in which information collected from various data sources are combined to help solve current issues. For example, following the terrorist events of September 11, 2001 in the United States, government programs sought to gather evidence and intelligence by combining information across various existing data collections [18]. While initial efforts succumbed to privacy concerns [5], on-going efforts to accomplish these goals continue.

To achieve the law-enforcement and counter-terrorism vision, two main hurdles must be overcome–one is a need for computational methods to combine disparate information accurately and the other is a need to sufficiently address privacy concerns.

"Link analysis" refers to a growing area of computer science that seeks to construct algorithms that learn information from disparate data [2]. Today, a fundamental motivation for link analysis development is law-enforcement and counter-terrorism. Example tasks include: name matching; link detection and social network analysis; detection and monitoring of intrusion, deception, conspiracy, fraud, and criminal activity; scene identification; person identification; and trend detection [14].

Because of the severity of harm to innocent individuals that can result, such uses immediately evoke privacy concerns (see [5] for an example). Examples of privacy concerns emerging from link analysis for law-enforcement and intelligence purposes include:

- the bulk of people whose information appears in these law-enforcement databases have done nothing to warrant suspicion.
- data captured in private spaces can be mixed with that obtained from public spaces, thereby thwarting protections afforded private spaces.
- individuals have no means to correct errors.
- no judicial review or impartial oversight exists to weigh societal benefits against individual risks in order to limit fishing expeditions and unwarranted inquiries.

These concerns are underscored by the government's mammoth power to take away an individual's liberty –i.e., to restrict a person's movement and autonomous self-determined behavior, in some cases indefinitely and without legal process [13] [17].[2]

2. PRIVACY BACKGROUND

Policymakers and computer scientists have previously addressed personal privacy issues in government databases. The dominant policy remedy allows individuals to review and correct personal information. Computer science remedies in other legal settings distort data such that resulting information remains useful while guaranteeing no one can be re-identified. Neither of these approaches may be best for the link analysis setting previously described. A claimed need for secrecy may override the policy remedy. An inability to identify beforehand which data elements are the valuable ones that should not be distorted makes prior computer science remedies difficult to consider. Detailed discussion on these issues appears in the next subsections.

[1] It is important to note that commercial applications of link analysis are also significant. These include insurance fraud detection, telecommunications network analysis, pharmaceuticals research, and epidemiology. While the focus of this paper is on homeland security uses, the issues and remedies presented in this paper are just as relevant to commercial practice though the motivations differ.

[2] Commercial harms to individuals include insurance coverage refusal, loss of credit worthiness, and denial of employment.

2.1 Fair Information Practices

Prior to the current era, there was a surge in government collection of information on individuals in the 1970's made possible by the growing availability of mini-computers. Privacy concerns voiced at that time culminated into a set of principles for privacy protection that have become known as the "Fair Information Practices." These principles form the basis of many policies and practices, most notably, the U.S. Privacy Act of 1974 and the European Union Directive on the Protection of Personal Data (1995). The basic principles are listed in Figure 1.

1	Existence of personal data collections should be public knowledge.
2	Individuals have a right to review and correct their information.
3	The minimum information necessary should be collected, and where appropriate, consent of the included individuals should be obtained.
4	Personal data should be accurate and complete and retained only for a given time period.
5	Data should only be used for the purpose originally intended.
6	Data should be protected by security safeguards against unauthorized access, modification or use.

Figure 1. Basic principles of the Fair Information Practices.

Some privacy advocates argue that Fair Information Practices should be imposed on information learned through link analysis [16]. Opponents argue that the Fair Information Practices are impractical for law-enforcement and counter-terrorism pursuits because potential criminals and terrorists cannot be given the opportunity to alter learned information or change behavior based on the knowledge of what has been learned.

2.2 Data Anonymity

Many regulations allow data to be shared beyond the original purpose of its collection, and without further adherence to Fair Information Practices, provided no one whose information is contained in the shared data can be re-identified. Examples include the U.S. medical privacy regulation known as HIPAA and Canadian and European data sharing practices. Data in which the subjects of the data can provably not be re-identified is termed "anonymous data."[3] Computer scientists have devised methods that guarantee a minimal risk that a subject of the data can be re-identified yet the data remains practically useful [9][7]. This is achieved by provably thwarting the ability to reliably link the anonymous information to other information that may lead to a re-identification. Therefore, anonymous data cannot be reliably linked to many kinds of data, thereby posing serious limitations on the ability to use anonymized data in link analysis.

[3] It has been shown that removing explicit identifiers, such as name, address, or Social Security numbers, or replacing them with made-up alternatives (no matter how strong the cryptographic hash) is not sufficient to render the result anonymous [5][2]. One way to provably anonymize data is k-anonymity [5], but a more real-world savvy approach is done using the Privacert Risk Assessment Server [12]. Many other frameworks are possible based on different statistical disclosure control techniques. On the other hand, algorithms emerging under the rubric of "privacy-preserving data mining" have not yet demonstrated their real-world applicability and legal appropriateness in the link analysis settings discussed in this paper.

One approach that might be useful when link analysis algorithms are deployed in the real world is Selective Revelation [9], which provides data to a surveillance system with a sliding scale of identifiability, where the level of anonymity matches scientific and evidentiary necessity. During normal operation, surveillance is conducted on sufficiently anonymous data that is provably useful. When sufficient and necessary scientific evidence merits, the system provides increasingly more identifiable data. 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. The limitation of its use in the link analysis setting previously described is that the role of particular data elements must be known beforehand, which is not always practical during the development of algorithms, but may be practical when deployed.

2.3 Computer Scientists to the Rescue

One effort that may help is to have the kinds of privacy protections provided in the Fair Information Practices be realized on results learned from link analysis. Doing so shifts the responsibility of privacy protection to the computer scientists who develop link analysis algorithms and to the experts who deploy them. This is the approach introduced in this paper.

In prior work [8], reactions by computer science researchers to privacy issues in their research was characterized by three positions: (1) "technology trumps privacy;" (2) "technology is policy neutral;" and, (3) "computer scientists take responsibility."

In the "technology trumps privacy" position, computer science researchers take stock in past accomplishments and computational benefits enjoyed by society, thereby relying on a belief that if society is forced to choose, it will choose advancements in computer technology over privacy. Warnings against this position caution that unforeseen dangers could be unleashed forever or the technology never deployed.

In the "technology is policy neutral" position, computer science researchers do not contemplate any privacy or social implications that may be inherent in the construction or existence of the technology they seek to build. Instead, these computer scientists want to pursue their research, leaving any related privacy issues to social scientists, policy makers, lawyers, and others. But some argue that such positions are themselves human value decisions, and computer science researchers cannot escape making them.

In the "computer scientists take responsibility" position, computer scientists take the initiative to incorporate privacy into their own constructions. Some believe assuming such responsibility is a necessary condition to insure viability of their technology. For those computer scientists, the next sections provide methods for incorporating privacy protections based on Fair Information Practices into newly constructed link analysis algorithms.

3. METHODS

For a link analysis algorithm to be put into practice in the settings previously described, the developers and/or those deploying the algorithm should provide a guarantee related to the utility of the algorithm (a "warranty") and a guarantee of privacy protections the algorithm provides (a "privacy statement"). The link analysis algorithm along with these accompanying guarantees describe the appropriate use of the algorithm; together, the algorithm and its guarantees are introduced as a "privacy-enhanced linking" solution. These are further described in the next subsections.

3.1 Privacy-Enhanced Linking

The term privacy-enhanced technology ("PET") has historically been used to generally refer to a technology that performs a task while providing privacy protection [11]. This writing extends the notion of PETs to privacy-enhanced linking (PEL) by dictating that one or more of the Fair Information Practices must be addressed within the link analysis algorithm and/or within the setting in which the algorithm is expected to execute. In PEL, two guarantees accompany the algorithm –a warranty statement and a privacy statement, as further described below.[4]

3.2 Warranty Statement

A PEL warranty statement addresses the quality of the algorithm as being suitable for, or adaptable to, a particular set of tasks. Computer scientists typically provide proofs of correctness and complexity when introducing an algorithm. These help characterize the utility that may be realized if the algorithm is put into actual practice, and therefore these will typically form the basis of a warranty statement for a link analysis algorithm.

3.3 Privacy Statement

Because it is believed that individuals cannot participate in link analysis settings sufficient to exercise Fair Information Practices, the onus of providing those protections shifts to the technology and is quantified and expressed in the PEL privacy statement.

Given a link analysis algorithm deemed appropriate for a particular setting, a PEL "privacy assessment" involves determining which Fair Information Practices are relevant and quantifying and characterizing algorithmic performance in terms of the protection provided. The results of the privacy assessment forms the basis for the PEL privacy statement.

The first principle found in the Fair Information Practices listed in Figure 1 may be beyond the scope of what can be accomplished by technical remedy, but the other principles, depending in part on the nature of the link analysis program used, can be realized by technology. A PEL privacy statement reports on the validity and interpretation of computed results as they relate to these principles. Here is an example.

Example. (The Watchlist Problem)

Government authorities have an explicit list of names of known or suspected terrorists (a "watchlist") they want to locate or merely track among the general population. There are vast numbers of locations the government seeks to query as to whether a person has appeared bearing the same explicit identity as one on the Watchlist. The idea is to review transactional data (store purchases, hotel registrations, airplane manifests, car rentals, school attendance records, etc) and match names to those on the watchlist. The problem is further complicated by the use of nicknames and misspellings [10].

[4] PEL is modeled after a new research paradigm (termed "unified computing") for constructing technology that is provably appropriate for a given setting. The developer provides warranty and compliance statements that show that the resulting technology remains useful while being compliant to the stated standard. (See privacy.cs.cmu.edu/dataprivacy/projects/unified/index.html for more information).

The principles of the Fair Information Practices that seem particular relevant are 2 and 4 in Figure 1. Because subjects of the data cannot review results learned from matching, it becomes extremely important that false positives (names of different people are incorrectly matched together) be rare. Preference should also be given to verified source information (e.g., from a credit card, driver's license) over casually acquired information.

The current solution involves the simple approach of matching names using soundex, which is a gross hash function in which spellings that may look or sound similar are hashed together [1]. Using soundex [4], the names "James" and "John" are hashed to J52 and J5, respectively, but the names "John," "Jane" and "Jean" are all hashed to the same "J5" value.

An accompanying PEL privacy statement would either include results of matching soundex names in a general population to report the false positives found or describe tests that should be conducted to determine whether the false positive rate for a given population is at an acceptable threshold.

By any reasonable standard on most large populations of names, soundex matching is not appropriate for this task, because it lumps too many different names together (see [1] for an example). Producing a PEL privacy statement revealed its inappropriateness.

Notice however, that the false negative rate (names for the same person incorrectly not matched together) is likely to be low (which is good), but this performance measure relates to the warranty, not the privacy statement.

4. EXPERIMENTAL RESULTS

An experiment was conducted to demonstrate the kinds of measurements that are likely to appear in PEL privacy assessments. The experiment involved automatically constructing a dossier on a given subject from information appearing on web pages indexed by Google. Information related to the subject was compiled into a single extended vita using semi-automated text extraction [12]. Human review was then conducted to assess the kinds of errors found.

4.1 Subject and Materials

The subject was Raj Reddy, a distinguished computer scientist and a Turing Award recipient. Entering "Raj Reddy" into Google generated 372,000 hits. The first 14,000 text pages were selected and the information surrounding the occurrence of his name was extracted and catalogued. Human review of the material was conducted. A few highlights showing ways linking can go wrong from a Fair Information Practices perspective are provided below.

4.2 False Positives and Negatives

There is a Raj Reddy, who is a reporter. Information associated with the reporter's activities were confused with those of Raj Reddy, the computer scientist. There were also situations in which "Reddy" did refer to Raj and other cases where it did not. There were a few cases referring to Helen Reddy, the singer.

4.3 Closed World Distortion

The only significant financial contribution found on-line was a $8000 gift to an organization. This experience provided a false over-emphasis of his enthusiasm towards this organization because his actual giving includes numerous gifts of larger amounts that were not listed on the web in any obvious manner.

4.4 Inflated Corroboration

There was more than 100 press articles that included his name, but many of them were simply variations (modified repeats) of fewer original articles. In fact the article having the most variation was neither the most insightful nor useful in learning facts about him. This experience warns that data may not reflect independent events.

4.5 Conflict Co-existence

Among the newspaper articles was one in which there was a quotation attributed to Raj harshly criticizing a computer company. Raj never knew of the existence of the article previously, and further he patently denies ever having made the comments attributed to him. This experience underscores a need to handle conflicts, assuming that in the absence of Raj's verbal input, the conflict would have been identified.

5. CONCLUSION

In conclusion, PEL provides a way for society to enjoy the benefits of link analysis while minimizing harm to individuals. The PEL privacy statement (a scientific assessment of the validity of results and of the appropriate use of the technology) does not actually provide privacy[5], but is consistent with minimizing the same kinds of harms as do the Fair Information Practices.

Support from the link analysis community is necessary if the potential of PEL is to be realized. First, publishing channels for the development of link analysis algorithms should include parts of papers or papers themselves that contain PEL privacy assessments, even if those assessments are critical or expose weaknesses in link analysis algorithms.

Second, the nature of PEL privacy assessments involves activities (e.g. testing the function on real-world data sets) that lie outside the kind of information normally included in computer science presentations of algorithms. These may rely on different scientific research methods (naturalistic observation, survey, interview, and experimentation) than traditional computer science research.

Third, computer scientists tend to exalt one algorithm over another if it solves more tasks. But PEL solutions are optimizations in which maximum utility is achieved while providing as much privacy protection as possible. PEL solutions modeling all relevant Fair Information Practices, while remaining useful, is most preferred.

6. ACKNOWLEDGMENTS

Special thanks to Lise Getoor and Chris Diehl for the opportunity to provide this paper. Much gratitude goes to Raj Reddy for recommending the experiment and for his willingness to be a consenting subject. This work was funded in part by a grant from the Intel Corporation. Additional support came from the Data Privacy Lab in the School of Computer Science at Carnegie Mellon University.

7. REFERENCES

[1] Goo, S. "Faulty 'No-Fly' System Detailed," Washington Post, October 9, 2004.

[2] Jensen, D. and Goldberg, H. (eds). *Artificial Intelligence and Link Analysis*. AAAI Fall Symposium. AAAI Technical Report FS-98-01. Orlando, Florida, October 23-25, 1998.

[3] Malin, B. and Sweeney, L. How (Not) to Protect Genomic Data Privacy in a Distributed Network: Using Trail Re-identification to Evaluate and Design Anonymity Protection Systems. *Journal of Biomedical Informatics*. 2004; 37(3): 179-192. (privacy.cs.cmu.edu/dataprivacy/projects/trails/dnaTrails.html)

[4] Russell, R. Soundex. U.S. Patent 1,261,167 April 2, 1918.

[5] Safire, W. "Dear DARPA Diary," New York Times, June 5, 2003.

[6] Sweeney, L. Information Explosion. *Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies*, L. Zayatz, P. Doyle, J. Theeuwes and J. Lane (eds), Urban Institute, Washington DC 2001. privacy.cs.cmu.edu/people/sweeney/explosion.html

[7] Sweeney, L. k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10 (5), 2002; 557-570. (privacy.cs.cmu.edu/people/sweeney/kanonymity.html)

[8] Sweeney, L. Navigating Computer Science Research Through Waves of Privacy Concerns: Discussions among Computer Scientists at Carnegie Mellon University, *ACM Computers and Society*. 34 (1), April 2004. (privacy.cs.cmu.edu/dataprivacy/projects/csresearch.html)

[9] Sweeney, L. Privacy-Preserving Surveillance using Databases from Daily Life. *IEEE Intelligent Systems*, 20 (5), September-October 2005. (privacy.cs.cmu.edu/dataprivacy/projects/selectiverevelation/index.html)

[10] Sweeney, L. Towards a Privacy-Preserving Watchlist. *AAAI Spring Symposium: AI Technologies for Homeleand Security*, 2005. (privacy.cs.cmu.edu/dataprivacy/projects/watchlist/index.html)

[11] 6th Workshop on Privacy-Enhancing Technologies. (petworkshop.org/2006/cfp.html)

[12] The AutoVita Project. (privacy.cs.cmu.edu/dataprivacy/projects/autovita/index.html)

[13] "Court Orders Due Process for Guantanamo Prisoners." Washington Times, June 29, 2004. www.washingtontimesh.com/national/20040629-121624-4193r.htm

[14] International Conference on Artificial Intelligence and Law. Workshop on Data Mining, Information Extraction, and Evidentiary Reasoning for Law Enforcement and Counter-Terrorism. Bolognam Italy. June 11, 2005. www.karlbranting.net/icail2005workshopcfp.html

[15] Privacert Risk Assessment Server. www.privacert.com

[16] Privacy Rules for Access to Personal Data. Center for Democracy and Technology. www.cdt.org/security/guidelines/

[17] U.S.A. Patriot Act, HR 3162. October 2001. thomas.loc.gov/cgi-bin/query/D?c107:4:./temp/~c107j6Mg7 b::

[18] "Your Papers Please," World Net Daily, Jan 16, 2003; www.worldnetdaily.com/news/article.asp?ARTICLE_ID=30523.

[5] Data anonymity does offer privacy, but link analysis algorithms many not work with anonymized data.