Privacy Technology for E-Governance

Jaijit Bhattacharya[1]

ABSTRACT

In providing services to the public and carrying out various functions, governments collect and use a wide range of personal information about their citizens. The introduction of e-government and the electronic delivery of services have further expanded government collection of personally-identifiable data. Thus, more and more government data is being put onto networks and are hence susceptible to violation of the privacy constraints of the data. The government agencies that collect personal information should adopt and maintain adequate privacy practices. In this paper, we outline these crucial privacy issues and present various solutions that are available.

1. Introduction

The fast progress in networking technologies has led to an enormous amount of digital information stored all over the world. This process has been accompanied by a rise of tools that are able to collect data, add them to databases and find information that could not be discovered in an obvious way. This explosive growth in digital data has brought increased concerns about the privacy of personal information. Privacy concerns restrict the free flow of information. Organizations do not want to reveal their private databases for various legal and commercial reasons. Neither do individuals want their data to be revealed to parties other than those they give permission to.

This is specially the case with E-Government, as it is an amalgam of interconnected heterogeneous information systems in which government agencies and public and private sectors exchange a high volume of information. Several government agencies have aggressively adopted information technologies in order to modernize the governments highly fragmented service-centric information infrastructure by improving information flow and the decision-making process. The E-Government infrastructure that essentially builds on these Internet technologies carries over a level of concerns for citizen privacy.

Citizens, while, welcoming client-driven, interactive, integrated information and services from the government, have concerns about privacy in electronic contexts. Numerous surveys (Westin, 1998; Cranor. et. al., 1999) have found that citizens will not participate in electronic transactions where privacy concerns have not been appropriately addressed.

In order to protect the concerns of citizens, many laws have been enacted to safeguard the privacy of their information. Laws alone cannot address all the concerns surrounding a complex issue like privacy. The total solution must combine policy, law, and technology. Therefore, adoption of new policies for collection and access, use, disclosure and retention of information, and for redress and oversight is vital. Moreover,

[1] Department of Management Studies, Indian Institute of Technology, Delhi, India, (E-mail : jaijit@dms.iitd.ernet.in, Telephone: +91)

technology itself should be a part of this solution. The same technology that permits the accumulation, sharing, and analysis of huge databases should also incorporate features that protect information from abuse or misuse into information sharing systems. In this paper, we outline these crucial privacy issues in E-Government systems and present various legislations and solutions that are available.

The paper is organized as follows. It begins with a discussion on privacy issues that exist in the E-Government systems. Section 3 describes the various legislations that have been enacted by the governments. Section 4 describes the various privacy technologies that are available.

2. Privacy Issues in E-Government

Governments are increasingly using the Internet as a means for the delivery of services and information. This development allows users to register for government services; obtain and file government forms; apply for employment; comment on public policy issues; and engage in a growing number of other functions - all on line. The trend towards E-Government and the electronic delivery of services has further expanded government collection of personally identifiable data. Governments' practices in collecting, retaining, and managing personal data pose a wide range of privacy concerns. With this increasing use of technology in government-to-citizen interactions, G2C, it is important to ensure that government agencies that collect personal information adopt and maintain adequate privacy practices.

Many details of an individual's life, activities and personal characteristics can be found scattered throughout the files of government agencies. Many of these records are, by law or tradition, open to public inspection. This transparency serves important democratic values. But in the Internet Age it also poses privacy risks. It is now increasingly possible to construct a detailed profile of individual using only publicly available, individually identifiable information from government records. While the types of government records that are publicly available vary from jurisdiction to jurisdiction, publicly accessible government records with personal information may include property ownership and tax records (name, address, value of property); driver's license (name, address, data on birth, physical characteristics, ID number); voter registration files; and occupational licenses.

In the Information Age, personal information has become a highly valued commodity that is collected, aggregated, shared and sold in ways never before imagined. Whole industries have formed solely to collect and distribute sensitive information that individuals once viewed as under their control: medical records, personal shopping habits and financial data. As public institutions move services on line, there is a growing risk of compromise and abuse. If personal identification data is used in the context of a given transaction, privacy concerns occur but seem manageable. However, privacy concerns become more serious when these data are the subject of secondary use by business and/or government. They arise because such use often means activities that do not reproduce social institutions in ways that allow continuation of the formally achieved, privacy-related status quo, but rather, transform social institutions in ways that force individuals to renegotiate arrangements for privacy protection in ways that diminish private space.

Secondary use of personal identification data does not stop at the boundary of the digital divide. The demand of markets for consumer-related information has been globalized. The demand of governments for information on the people may be local, but it is increasingly supported by Information and Communication Technology applications on both sides of the digital divide. The growing experience of people adversely affected by the secondary use of personal identification data limits the element of trust in transactions that require revealing such data. Within the context of e-government applications in particular, the willingness of many to reveal personal information may be marred by the lack of concern on the part of the authorities for protecting it or clearly indicating all other purposes for which it may be used. As a result, such information is given if the ends to be achieved are worth the price of potentially diminished privacy,

e.g. in the context of welfare programmes.

3. Privacy Legislations

It is gradually being accepted that privacy is an important right that the state has some obligation to protect through regulatory policy. Privacy laws embody the premise of trust and confidence between the citizenry and the government when it comes to the delivery mechanisms and ingrained e-government programs. Privacy is the right to have an individual's personal information protected from the undue prying eyes of governments and of organizations seeking to use such personal information for trade and profit, without the consent of the individual, except in exceptional circumstances dictated by law. The new rules laid out for people to have access to their own personal data represent a transfer of power from the state to the citizen in that the citizen now has some control over personal information.

At the beginning of the computer revolution, governments developed a set of Principles of Fair Information Practices. These principles are intended to foster individuals' control over their personal information, limit data collection and place responsibilities on data collectors. They are the basis for most modern data protection and online privacy laws and policies. Many countries have adopted national privacy or data protection laws. Such laws may apply to data about individuals collected by the government, to personal data in the hands of private sector businesses, or to both.

According to US Privacy Act of 1974 (US Privacy Act of 1974), a privacy sensitive transaction will permit:

- An individual to
  - determine what records pertaining to him are collected, maintained, used, or disseminated
  - prevent records pertaining to him, obtained for a particular purpose from being used or made available for another purpose without his consent
  - gain access to information pertaining to him in records, and to correct or amend such records
- Collect, maintain, use or disseminate any record of personal identifiable information in a manner that assures that:
  - such action is for a necessary and lawful purpose,
  - that the information is current and accurate for its intended use,
  - that adequate safeguards are provided to prevent misuse of such information
- Permit exemptions from the requirements with respect to the records provided in this act only in those cases where there is an important public policy need for such exemption as has been determined by specific statutory authority
- Be subject to civil suit for any damages, which occur as a result of willful or intentional action, which violates any individual's right under this act.

Similarly, there are other governments and other related US acts and those of the European Union (U.S.Federal Trade Commission report to Congress, May 2000; US Computer Matching and Privacy Protection Act of 1998; Official Journal of the European Communities, 1995; World Wide Web Consortium, Platform for Privacy Preferences 1.0 (P3P 1.0)Specification,W3C Working Draft, 2000.) that define the Privacy rights of individuals.

The Freedom of Information and Protection of Privacy Act (FIPPA) (Freedom of Information and Protection of Privacy Regulations) provides a right of access to records of public bodies, subject to certain specified exceptions, and with protection for personal information held by public bodies.) The purposes of this Act are to make public bodies more accountable to the public and to protect personal privacy by

- giving the public a right of access to records,
- giving individuals a right of access to, and a right to request correction of, personal information about themselves,

- specifying limited exceptions to the rights of access,
- preventing the unauthorized collection, use or disclosure of personal information by public bodies, and
- providing for an independent review of decisions made under this Act.

Also, in order to tackle the growing problem of identity theft, Senator Feinstein introduced Package of Bills. Senator Feinstein's package of bills includes:

- The Privacy Act

A comprehensive bill that would set a national standard for protecting personal information such as Social Security numbers, driver's licenses, and medical and financial data, including information collected both online and offline. Modeled on California's financial privacy law, it requires companies to let consumers "opt in" before their most sensitive information is shared.

- The Social Security Number Misuse Prevention Act

This bill would regulate the use of Social Security numbers by government agencies and private companies by prohibiting the sale or display of Social Security numbers to the general public and by requiring Social Security numbers to be taken off of public records published on the Internet.

- The Notification of Risk to Personal Data Act

Modeled on California's database security law, this bill would define as personal data an individual's Social Security number, driver's license number, state identification number, bank account number or credit card number; require a business or government entity to notify an individual when it appears that a hacker has obtained unencrypted personal data; levy fines by the FTC of $5,000 per violation or up to $25,000 per day while the violation persists; and allow California's privacy law to remain in effect, but preempt conflicting state laws.

In addition to the above provisions of the various privacy acts, a privacy-sensitive transaction need to also permit an individual to securely transact without revealing his or her identity and ensure that the transaction is upheld in the court of law.

4. Privacy Enhancing Technologies

An essential aspect of any privacy protection regime is enforcement. Without privacy policy enforceability across enterprises, everything is a matter of trust. Therefore, it is recommended to use privacy-enhancing technologies in e-governments as a natural part of the development of an Internal Market (the free flow of personal information) and the protection of the fundamental rights and freedoms of individuals. A number of technologies exist that can help government to effectively address privacy issues in any E-Government initiative. Primary among them has been to (a) define a privacy specification language, (b) enforcing privacy during data mining, (c) ensuring that the database itself ensures privacy (Agrawal et. al., 2002), (d) ensuring transactional Privacy using encryption and co-processors , (e) Statistical Disclosure Control, (f) Anonymized data analysis and (g) developing a privacy broker for privacy preserving transactions.

4.1 Privacy Specification Language

The E-Government Act of 2002 requires federal agencies to put in place privacy protections for information collected electronically. Specifically, Section 208 outlines requirements for privacy policies on federal government Web sites that collect information. As a result of this legislation, privacy policies in standardized machine readable formats should be available on all agency Web sites.

A World Wide Web Consortium standard, the Platform for Privacy Preferences, or P3P (http://www. w3.org/P3P/P3FAQ. html.), is a broadly adopted formal language for communicating privacy promises to consumers. A P3P policy is a promise by a service provider to limit the use of certain data to certain purposes, recipients, and retention periods. Prior to retrieving a web page, a consumer's web browser first

downloads the site's P3P policy, and then compares the downloaded policy against its user's privacy preferences. If the policy respects the user's preferences, the web browser retrieves the web page. However, if the policy does not respect the user's preferences, the browser may block the site or notify the user. When manipulating data, the web site operator is obligated to adhere to the P3P policy under which it collected the data.

However, P3P does not provide an enforcement mechanism for organizations to use in monitoring their information handling practices. IBM's Enterprise Privacy Authorization Language (EPAL) (Ashley, et. al. 2003) addresses the need for machine enforceable policies. Like P3P, EPAL is an XML-based privacy policy specification language and is designed for organizations to specify internal privacy policies. These EPAL policies can be used internally and amongst the organization and its business partners to ensure compliance with the underlying policies of each.

Posting privacy policies is essential in building trust between government websites and their users and these policies are created to inform users of a site's data collection, use and disclosure practices.P3P is not a panacea for privacy, but it does represent an important opportunity to make progress in building greater privacy protections into the web experience of the average user.

4.2 Enforcing privacy during data mining

Privacy is becoming an increasingly important issue in counter-terrorism and homeland defense-related applications. These applications may require creating profiles, constructing social network models, and detecting terrorist communications among others from privacy sensitive data.

One method for preserving individual's privacy is by distorting the data values. The idea is that the distorted data does not reveal private information, and thus is "safe" to use for mining. The key result is that the distorted data, and information on the distribution of the random data used to distort the data, can be used to generate an approximation to the original data distribution, without revealing the original data values. Consider the example of census data: the government of a country collects private information about its inhabitants, and then has to turn this data into a tool for research and economic planning. However, it is assumed that private records of any given person should not be released nor be recoverable from what is released. In particular, a company should not be able to match up records in the publicly released database with the corresponding records in the company's own database of its customers. Therefore, distortion can be used to ensure high privacy protection.

Agrawal and Srikant first proposed using randomization to solve the above problem (Agrawal and Srikant, 2004). In their randomization scheme, a random number is added to the value of a sensitive attribute. For example, if xi is the value of a sensitive attribute, xi+r, rather than xi, will appear in the database, where r is a random value drawn from some distribution. It is shown that given the distribution of random noises, recovering the distribution of the original data is possible. The randomization techniques have been used for a variety of privacy preserving data mining work (Agrawal and Aggarwal, 2001; Rizvi and Haritsa, 2002; (Du and Zhan, 2003).

Another approach to achieve Privacy-Preserving Data Mining is to use Secure Multi-party Computation (SMC) techniques. SMC deals with computing certain function on multiple inputs, in a distributed network where each participant holds one of the inputs; SMC ensures that no more information is revealed to a participant in the computation than what can be inferred from the participant's input and the final output. For example a government agency might have employment information, another health data, and third information about education. An analysis on an integrated database would be more informative and powerful than, or at least complementary to, individual analyses.

The work in (Du and Attalah, 2001) proposes a transformation framework that allows to systematically transforming normal computations to secure multiparty computations. Among other information items, a discussion on transformation of various data mining problems to a secure multiparty computation is demonstrated. These problems include privacy preserving information retrieval, privacy preserving geometric computation, privacy preserving statistical analysis, and privacy preserving scientific computations, etc.

The work in (Clifton, et. al.,2002.) discussed several SMC protocols to securely compute sum, set union, size of set intersection and inner product. These protocols are directly applied for privacy preserving association rule mining from vertically partitioned data and horizontally partitioned data, clustering with distributed EM mixture modeling, and K-Means clustering over vertically partitioned data. The SMC ideas have also been applied for privacy preserving distributed decision tree induction, naive Bayes classification for horizontally partitioned data, privacy preserving Bayesian network structure computation for vertically partitioned data , and many else.

4.3 Privacy Preserving Databases

Oracle has implemented privacy (Edwards) using a combination of techniques that allow a higher granularity of control at tuple level as well as at column level. The key mechanisms are as follows:

- Strong authentication and single sign-on: Strong authentication is generated by PKI infrastructure that uses industry standard X.509 digital certificates for strong authentication
- Granular Access control through views: A view is a subset of one or more tables. However, views have issues of scalability and complication in administration of security and privacy.
- Virtual Private Database (row level control): VPD enables, within a single database, per user or per group data access with the assurance of data separation. By dynamically appending SQL statements with a predicate (a "where" clause), VPD limits access to data at the row level and ties security policy to the table itself.
- Label-Based Access Control: The label security mediates access to data by comparing a sensitivity label on a piece of data with label authorizations assigned to an application user. Such access mediation allows data to be separated into different sensitivities within a single database.
- Secure Application Role: It ensures that the appropriate conditions are met before the user can excursive privileges granted to the role in the database. This limits the bypassing of the application to directly access the database.
- Encryption in the database: Oracle supports DES (56 bit) and triple DES (112 and 168 bits) encryption of the records.

However, Oracle 9i's solution is not a tool dedicated for privacy but it is a tool that facilitates privacy-enabled implementations.

Another approach is that of Hippocratic databases (Rizvi and Haritsa, 2002) that uses components of secure database and introduces privacy control within the database itself. The Hippocratic database uses Privacy Metadata, which is defined as (a) External recipients, (b) Retention period and (c) Authorized users.
In providing services to the public and carrying out various functions, government collects and uses a wide range of personal information about the people. Different individuals will have different choices pertaining to sharing of their personal information. Government can deploy a Hippocratic database to support the privacy needs of individuals.

5. Ensuring Transactional Privacy using Encryption and co-processors

With the use of E-Government systems, database transactions are executed over the internet and the data is accessible through the web. One method to secure this data is to encrypt it. If hacked into, the hackers

would only get a string of garbage and nothing meaningful. The only people who could use the data would be those individuals having the encryption key.

An uncompromised program (e.g. IBM 4758 programmable secure coprocessor) as a broker for all database transactions can be used. The uncompromised program encrypts the stored data with its private key and signs the outgoing data with its private key again (Kaplam, 1996; Smith and Safford, 2001; Smith, 2000). Alternatively, privacy of data collection is ensured by using a direct encrypted connection between the database and the user's client (Oracle Corporation. Database Encryption in Oracle 8i, 2000).

The work in (Mattsson) presents a scalable approach for data privacy and security in which a security administrator protecting privacy at the level of individual fields and records, and providing seamless mechanisms to create, store, and securely access databases. Such a model alleviates the need for organizations to purchase expensive hardware, deal with software modifications, and hire professionals for encryption key management development tasks. They proposed, implemented, and evaluated different encryption schemes.

5.1 Statistical Disclosure Control

Privacy in statistical databases, known as Statistical Disclosure Control (SDC), seeks to protect statistical data in such a way that they can be publicly released without giving away confidential information that can be linked to specific individuals or entities. Most countries have legislation which compels national statistical agencies to guarantee statistical confidentiality when they release data collected from citizens or companies.

The problem of protecting sensitive information in a database while allowing statistical queries (i.e. queries about sums of entries, and the like) has been studied extensively since the late 70's. In their comparative survey of privacy methods for statistical databases, Adam and Wortmann (Adam and Wortmann, 2004) classified the approaches taken into three main categories: (i) query restriction, (ii) data perturbation, and (iii) output perturbation.

Query Restriction: In the query restriction approach, queries are required to obey a special structure, supposedly to prevent the querying adversary from gaining too much information about specific database entries. The limit of this approach is that it allows for a relatively small number of queries.

Data/Output Perturbation: In the data perturbation approach queries are answered according to a perturbed database. In the output perturbation approach, the database first computes an 'exact' answer, but returns a 'noisy' version of it. Methods of data perturbation include swapping where portions of the data are replaced with data taken from the same distribution, and fixed perturbations where a random perturbation is added to every data entry. Methods of output perturbation include varying output perturbations, where a random perturbation is added to the query answer, with increasing variance as the query is repeated, and rounding either deterministic or probabilistic.

The three sub disciplines of Statistical Disclosure Control are:
Tabular data protection: This is the oldest and best established part of SDC, because tabular data have been the traditional output of national statistical offices. The goal here is to publish static aggregate information, i.e. tables, in such a way that no confidential information on specific individuals among those to which the table refers can be inferred (Adam and Wortmann, 2004; Willenborg and DeWaal, 2001).

Dynamic databases: The scenario here is a database to which the user can submit statistical queries (sums, averages, etc.). The aggregate information obtained by a user as a result of successive queries should not allow him to infer information on specific individuals. Since the 80s, this has been known to be a difficult

problem, subject to the tracker attack [85]. One possible strategy is to perturb the answers to queries; solutions based on perturbation can be found in (Duncan and Mukherjee. 2000) If perturbation is not acceptable and exact answers are needed, it may become necessary to refuse answers to certain queries; solutions based on query restriction can be found in (Chin and Ozsoyoglu, 1982) and (Gopal, et. al. 1998). Finally, a third strategy is to provide correct (unperturbed) interval answers, as done in (Garfinkel, 2004) and (Gopal et.al.2002).

Microdata protection: This subdiscipline is about protecting static individual data, also called microdata. It is only recently that data collectors (statistical agencies and the like) have been persuaded to publish microdata. Therefore, microdata protection is the youngest subdiscipline and is experiencing continuous evolution in the last years (Crises, 2004; Gopal, et. al. 1998).

5.2 Anonymized Data Analysis

In order to make progress in improving the nation's response to terrorism and preserving civil liberties, the government uses commercial and governmental databases to collect information about individuals. When personally identifiable information is used to make judgments about people, a person sometimes will be misidentified as a criminal or a suspected terrorist or a risk when in fact he is innocent but shares some identifiers with someone who is of interest to the government.

Anonymizing technology would allow multiple data holders to collaborate to analyze information while protecting the privacy and security of the information. If both the privacy of personal information and the operational sensitivity of the information the government has on known or suspected terrorists can be assured, the reluctance to share data would be minimized. This would enable analysis of data from diverse sources, without requiring data to be gathered in a single place in a form that could be read or used for other purposes. This would limit abuses, including mistaken identity. The process of -anonymizing a dataset involves applying operations to the input dataset including data suppression and cell value generalization. Suppression is the process of deleting cell values or entire tuples. Generalization involves replacing specific values such as a phone number with a more general one, such as the area code alone.

There are several -anonymization algorithm proposals in the literature. Iyengar (Iyengar, 2002) shows how to attack a very flexible (and highly combinatorial) formulation of -anonymity using a genetic algorithm. The datafly approach of Sweeney (Sweeney, 2002) is another greedy approach that generates frequency lists and iteratively generalizes those combinations with less than occurrences. Like incomplete stochastic approaches, iterative greedy approaches such as -argus and Datafly offer no solution quality guarantees. Sweeney (Sweeney, 2002) and Samarati (Meyerson and Williams, 2004) have both proposed complete algorithms for -anonymization. Sweeney's algorithm exhaustively examines all potential generalizations to identify the optimal (or "preferred") generalization that minimally satisfies the anonymity requirement, acknowledging the approach is impractical even on modest sized datasets. Samarati proposes an algorithm to identify all "k-minimal" generalizations, among which reside the optimal k-anonymizations according to certain preference criteria.

5.3 Privacy Broker for Privacy Preserving Transactions

The Privacy Broker (Bhattacharya and Gupta, 2004) for privacy preserving transactions enables the following aspects of privacy without modifying the database kernel. (a) The Broker accepts the agreed privacy specification and ensures adherence of the stated privacy policies, (b) it enables individuals to authorize specific individuals to access their data and (c) it enforces non-repudiation of agreements between visitors and web-sites.

The overall structure of the broker is as follows:

- Uses an uncompromised program as a broker for all database transactions

- The uncompromised program encrypts the stored data with its private key and signs the outgoing data with its private key again
- All data accesses are through "Capability Certificates"
- "Capability Certificates" also double up to support non-repudiation
- Capability certificates will allow a suitably authorized person to allow a user to access privacy constrained data.
- Such authorization would be through capability certificates, which would allow the user to access data for a pre-specified time period. For example, the Mayor of a city can allow the local police head to access medical data, for forensic reasons, of all citizens in the city who have blue eyes and are of the age between 20 and 30 years and NO other data.
- The capability certificates would allow the appropriate policy to be executed, fetching the required data

To capture privacy policies for managing databases this approach uses a layered architecture for a policy based data administrator (Batra, et. al., 2002). The policies are defined by the decision makers/ data administrator using a friendly graphical user interface and then these policies are modeled as ECA (Event-Condition-Action) like rules.

This Privacy Broker can be easily used in any E-Government initiative to reduce privacy violation risk and enforce the committed privacy policies. Such a Broker-based approach ensures that the solution is independent of the database used. It also facilitates privacy administrators, with low IT skills, by setting privacy policies for managing the system.

6. Concluding Remarks

The concern among citizens about how their personal data is stored, processed and transmitted in an e-government context is among the top e-government barriers. Privacy laws embody the premise of trust and confidence between the citizenry and the government. However, privacy protection can only be guaranteed through the laws of mathematics rather than the laws of men and whims of bureaucrats. Therefore, E-Government systems should incorporate privacy enforcement mechanism to enforce privacy rights for citizens that are enshrined in the laws.

References

1. A.F. Westin. E-commerce and privacy: What net users want. Technical report, Louis Harris & Associates, June 1998. Available from http://www.privaeyexehange.org/iss/surveys/eeommsum, html.
2. L.F. Cranor, J. Reagle, and M.S. Ackerman. Beyond concern: Understanding net users' attitudes about online privacy. Technical Report TR 99.4.3, AT&T Labs-Research, April 1999. Available from http://www.research.att.corn/1ibrary/trs/TRs/99/99.4/99.4.3/report.htm.
3. The World Wide Web Consortium. The Platform for Privacy Preference (P3P}. Available from http ://www. w3.org/P3P/P3FAQ. html.
4. US Privacy Act of 1974
5. U.S.Federal Trade Commission report to Congress, May 2000. Privacy Online: Fair Information Practices in the Electronic Marketplace.
6. US Computer Matching and Privacy Protection Act of 1998
7. Official Journal of the European Communities of 23 November 1995 No L.281 p.31.Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
8. World Wide Web Consortium, Platform for Privacy Preferences 1.0 (P3P 1.0)Specification,W3C Working Draft 10 May 2000.
9. R. Agrawal, J. Kiernan, R. Srikant, Y. Xiu. *Hippocratic Databases (Vision Paper)*. IBM Almaden Research Center. 2002
10. Marc A. Kaplam. IBM Cryptolopes TM, Superdistribution and Digital Rights Management.

http://www.research.ibm.com/people/k/kaplan/cryptolope-docs/crypap.html, Dec 1996
11. S.W.Smith, D. Safford. Practical Server Privacy with Secure Coprocessors. In IBM Systems Journal, Vol 40, No.3, 2001
12. S.W. Smith, WebALPS: Using Trusted Co-Servers to Enhance Privacy and Security of Web Interactions, Research Report RC-21851, IBM Thomas J. Watson Resrach Center, Yorktown Heights, NY 10598 (October 2000).
13. Oracle Corporation. Database Encryption in Oracle 8i, August 2000.
14. Batra, V.; J. Bhattacharya; H. Chauhan; A. Gupta; M. Mohania; U. Sharma. 2002. "Policy Driven Data Administration". In POLICY 2002, IEEE 3rd International Workshop on Policies for Distributed Systems and Networks
15. D. Agrawal and C. C. Aggarwal. On the design and quantification of privacy preserving data mining algorithms. In Proceedings of the Twentieth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, Santa Barbara, California, USA, May 21-23 2001. ACM.
16. K.B. Edwards. Oracle 9i Privacy Protections, Oracle Corporation
17. Bhattacharya J. and Gupta S.K., 'Privacy Broker for Enforcing Privacy Policies in Databases', KBCS-2004. Fifth international conference on knowledge based computer systems. Hyderabad, India, December 19-22, 2004
18. Rakesh Agrawal and Ramakrishnan Srikant. Privacy-preserving data mining. In Proceedings of the ACM SIGMOD Conference on Management of Data,Pages 439–450, 2000.
19. Shariq J. Rizvi and Jayant R. Haritsa. Maintaing data privacy in association rule mining, In Proceedings of the 28th International Conference on Very Large Databases, 2002.
20. W. Du and Z. Zhan. Using Randomized Response Techniques for Privacy-Preserving Data Mining. In Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Washington, DC, USA.. pages-.505-510, 2003.
21. Wenliang Du and Mikhail J. Attalah, Secure multi-problem computation problems and their applications: A review and open problems, TechReport CERIAS Tech Report 2001-51, Center for Education and Research in Information Assurance and Security and Department of Computer Sciences, Purdue University, West Lafayette, IN 47906, 2001.
22. Chris Clifton, Murat Kantarcioglou, Xiadong Lin, and Michael Y. Zhu, Tools for privacy preserving distributed data mining, In Proceeding of SIGKDD Explorations, no. 2, 2002.
23. "Freedom of Information and Protection of Privacy Regulations".
24. P.Ashley, S. Hada, G.Karjoth, C. Powers and M. Schunter, Enterprise Privacy Authorization Language 1.2 (EPAL) W3C Member Submission, November 2003
25. Mattsson, Ulf T., "Providing Database Encryption as a Scalable Enterprise Infrastructure Service" . http://ssrn.com/abstract=664201
26. N. R. Adam and J. C. Wortmann, Security-Control Methods for Statistical Databases: A Comparative Study, ACM Computing Surveys 21(4): 515-556 (1989). S. Giessing. Survey on methods for tabular data protection in argus. In J. Domingo-Ferrer and V. Torra, editors, Privacy in Statistical Databases, volume 3050 of LNCS, pages 1–13, Berlin Heidelberg, 2004. Springer.
27. L. Willenborg and T. DeWaal. Elements of Statistical Disclosure Control. Springer-Verlag, New York, 2001.
28. F. Y. Chin and G. Ozsoyoglu. Auditing and inference control in statistical databases. IEEE Transactions on Software Engineering, SE-8:574–582, 1982.
29. R. Gopal, R. Garfinkel, and P. Goes. Confidentiality via camouflage: the cvc approach to disclosure limitation when answering queries to databases. Operations Research, 50:501–516, 2002.
30. R. Gopal, P. Goes, and R. Garfinkel. Interval protection of confidential information in a database. INFORMS Journal on Computing, 10:309–322, 1998.
31. G. T. Duncan and S. Mukherjee. Optimal disclosure limitation strategy in statistical databases: deterring tracker attacks through additive noise. Journal of the American Statistical Association, 95:720–729, 2000.
32. R. Garfinkel, R. Gopal, and D. Rice. New approaches to disclosure limitation while answering queries to a database: protecting numerical confidential data against insider threat based on data and algorithms, 2004. Manuscript. Available at http://www.eio.upc.es/seminar/04/garfinkel.pdf.
33. G. Crises, Microdata Disclosure Risk in Database Privacy Protection, Research Report CRIREP-04-003, Sep. 2004.
34. G. Crises, An Introduction to Microdata Protection for Database Privacy, Research Report CRIREP-04-006, Sep. 2004.

35. V. Iyengar. Transforming data to satisfy privacy constraints. In Proc. of the Eigth ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining, 279-288, 2002.
36. A. Meyerson and R. Williams. On the complexity of optimal k-anonymity In Proc.of the 23rd ACM SIGMOD-SIGACT-SIGART Symposium on the Principles of Database Systems, 223-228, 2004
37. L. Sweeney. Achieving k-anonymity privacy protection using generalization and suppression. Int'l Journal on Uncertainty, Fuzziness, and Knowledge-Base Systems 10(5): 571-588, 2002.

About the Author

Jaijit Bhattacharya is Adjunct Faculty in the Department. He also serves as Country Director, Government Strategy at Sun Microsystems. He is responsible for the creation of the next generation of solutions for governments, based on open standards. He is an e-Governance advisor to Government of Sri Lanka and has been conducting training for ADB institute in Tokyo on Public Expenditure Management and has helped World Bank develop curriculum for their e-Leadership program. He has also been delivering lectures at INSEAD Singapore campus and at IIM Calcutta. Dr. Bhattacharya has developed business models and strategies for leading companies in the IT, media and computer hardware industries. He has numerous research papers to his credit in leading journals and conferences. He has edited two book "Technology in Government", "Towards Next Generation of E-government" and has co-authored 'Government On-line-Opportunities and Challenges', published by Tata McGraw Hill which was released by the Honourable President of India, Shri Abdul Kalam. Dr. Bhattacharya did his PhD from Department of Computer Science and Engineering IIT Delhi, MBA from Indian Institute of Management, Calcutta and B.Tech in Electrical Engineering from Indian Institute of Technology, Kanpur.