Fundamental Limits on a Model of Privacy-Trust Tradeoff: Information Theoretic Approach

Garimella Rama Murthy

Department of Computer Science and Information Engineering
International Institute of Information Technology—Hyderabad
Gachibowli, HYDERABAD-500032, AP, India. (Email: rammurthy@iiit.net)

(Received Aug. 21, 2005; revised and accepted Oct. 11, 2005)

Abstract

Zhong et al. formulated the problem of disclosing credentials (associated with privacy) for building trust in an open environment like Internet. Thus, one form of privacy-trust problem is formulated in [17]. In [17], an entropy measure is defined to quantify privacy loss. In this research paper, by a proper formulation (modelling), the privacy-trust tradeoff is attacked from the Information Theoretic view point. It is shown that as long as privacy loss (or more accurately a measure related to privacy loss) is below a limit, through the use of properly coded credentials, it is possible to achieve arbitrarily good trust. The approach presented here will lead to several practical schemes in trading privacy for trust.

Keywords: Digital credentials, encoding, privacy, trust

1 Introduction

The dawn of information age with the progress of hardware technology has made it easy to store and process large amounts of transactional information. Data mining is one such area which received increasing attention in the recent years. It is considered a challenge to privacy preservation due to their natural tendency to use sensitive information about individuals. In the survey conducted by Forrester Research [11], it is reported that online consumers seriously worry about the personal information they divulge online. The report also declares that the fear of privacy protection held back roughly $15 billion in e-commerce revenue in 2001. On the contrary, incentive and monitoring mechanisms [2], including reputation systems [4, 10], suggest that a highly trusted user can get more benefits.

Digital credentials, such as certificates [5], recommendations or past transaction history [6] are utilized to build trust in an open environment like the Internet. But revealing these credentials may lead to compromise of privacy (to some degree) in the form of revelation of identity, shopping preferences etc. Thus, in such environment, **privacy and trust are in an adversarial relationship**. Thus giving up only the degree of privacy absolutely necessary to gain a level of trust is an important research problem. This problem requires quantification of tradeoff between privacy and trust. This tradeoff can be further decomposed as 3 sub-problems:

- Quantification of privacy lost by a specific piece of information.
- Quantification of benefit gained by having higher level of trust.
- Quantification of privacy sacrificed for certain amount of trust gain.

Thus intelligent privacy-for-trust decision can be made through a solution of the trade-off. Some research directions in solution of tradeoff are summarized in [17]: This research paper is organized as follows. In Section 2, survey of related literature is reported. Modelling of the privacy-trust trade-off problem is discussed in Section 3. In Section 4, basic limitations on a model of privacy-trust tradeoff are discussed. This section highlights the information theoretic approach to the privacy-trust tradeoff problem. Practical implementation issues are discussed in Section 5. The research paper concludes in Section 6.

2 Survey of Related Literature

Collecting pieces of information from different sources and putting them together to reveal private information is termed as "data fusion" [14]. From the data exchanged between applications, the ability to filter identifying information defines "data privacy". The growth of information gathered on individuals as well as organizations is leading to invasive data fusion. Consequently an entity may become increasingly reluctant to disclose private information [17].

Various researchers attempted to quantify privacy using different metrics. For instance, Rieter and Rubin [9]

use the size of the anonymity set (all the potential subjects that might have sent/ received data) to measure the degree of anonymity. In this approach the authors assume that each sender in the set has an equal probability of sending a message. Serjantove et-al [13] use entropy to measure the anonymity a system can achieve. In the context of design of privacy preserving data mining algorithms, Agarwal et al [1] use differential entropy to quantify the closeness of an attribute value estimated by an adversarial to its original value. This approach is similar to the query independent privacy loss definition in [17].

The issue of iteratively exchanging credentials between two entities to incrementally establish trust is investigated in the research on "AUTOMATED TRUST NEGOTIATION" [16]. In this approach the tradeoff between length of negotiation, amount of information disclosed and computational effort is considered. In [17], entropy is used to quantify privacy loss. Trust lifecycle management is proposed in [15] leading to trust based decision making. Also in [3] trust and evidence formalization is reported. In [12] Seigneur and Jensen propose an approach to trade minimal privacy for the required trust.

3 Modelling of Privacy-Trust Tradeoff

There are several ways of trading privacy for building trust in an environment like a Local Area Network (LAN) or a Wide Area Network (WAN) such as Internet. One of the most practical approaches is to issue digital credentials. These credentials are utilized for trading privacy for trust. In trading privacy, the following assumptions hold true.
Assumptions:

- User has multiple choices on what information to disclose.
- Each user can make his/her decision independently.

Let the private attributes that we want to conceal be $a_1, a_2, \cdots, a_m$. Also let the set of credentials a user has be $\{ c_1, c_2, \cdots, c_m\}$. Each attribute is encoded using a set of credentials (as discussed in Section 4). The encoding procedure is only partially known to the user who presents the credentials (to estimate the attribute value). Thus there will be an error in estimating the attribute value.

To introduce complexity into the privacy-trust tradeoff problem; we consider query-dependent and query independent privacy loss.

Estimation of Privacy Loss:

Entropy is a concept which originated in statistical mechanics. It measures the randomness in a system. Shannon utilized the same name to arrive at uncertainity associated with a random variable (Modelling the information transmitted by a source). Building on this notion, Zhang and Bhargava used entropy concept to measure the privacy loss for disclosing a credential. In their approach, the idea is that when an adversarial gains more credentials, the uncertainty about subjects is decreased. It should be noted that the author's approach of quantifying concepts like "privacy" and "trust" is not quite the same as that of Zhang et al.

In the following, we discuss two methods for evaluating query-dependent and query-independent privacy losses, respectively. The definitions from [17] are repeated here for the sake of clarity.

4 Basic Limitations on Privacy-Trust Tradeoff: Information Theoretic Approach

The motivation for the present research paper is shown in the following:

- As "measures related to privacy" is decreased, trust keeps increasing. It may be that after a while "measure related to privacy" is completely compromised and trust reaches its limit. Thus we are naturally led to the following question.

Question:

Under reasonable modelling assumptions, is it possible to show that as long as a "measure related to privacy" is below certain threshold, arbitrarily good value of "measure related to trust" is achieved.

The chief contribution of this paper is to invoke the results from information theory to derive fundamental limits on privacy-trust tradeoff (under practical modelling assumptions).

It should be noted that Agrawal et al. propose a metric for privacy based on Entropy in [1]. The metric used in this paper is easily related to the one in [1] using a monotone transformation. Also Serjantove et al. propose an information theoretic metric for anonymity in [13].

Clear Statement of the Problem Addressed:

As discussed in the previous section, to build trust in estimating the attribute of interest (say "age"), some privacy will be lost. This privacy loss is effected by issuing credentials. For the sake of simplicity, the credentials assume only binary values (generalization to the case of non-binary credentials is straightforward).

An attribute value is coded using credentials which assume binary values. The number of credentials utilized to code an attribute is a measure related to privacy. This leads to the so called "information vector" (in the terminology of information theory). It is intuitively clear that by adding arbitrarily large amount of redundancy (i.e. by padding "coding bits" to the information vector), the **probability of error in estimation of attribute value** (a measure related to trust) can (by the receiver)

be made arbitrarily small. Thus in such a situation, the data rate is becoming very small. Equivalently, a measure related to privacy loss is becoming arbitrarily large (by utilizing an arbitrarily large number of credentials for coding the attribute).

The question of interest is Q: Is it possible to show that as long as the **data rate (a measure related to privacy compromise)** is below a finite limit (but not zero) determined by the channel (associated with the attribute communication mechanism); is it possible to achieve arbitrarily high TRUST/RELIABILITY in estimating the attribute by the receiver (i.e. probability of error in estimation is arbitrarily small).

In order to invoke the ideas/concepts from information theory, the definitions on privacy loss are interpreted, modified in the following manner.

Modelling Assumptions: (Query-Independent Privacy Loss)

- Let a source of information be holding an attribute "a_j" that has a finite domain $\{v_1, v_2, \cdots, v_k\}$.
- For the sake of concreteness, let an attribute $\{a_j\}$ like 'age' take on 16 values with some discrete probability distribution. Let these attribute values be coded using 4 bits.
- Let the **credentials** utilized for encoding attribute values take on "binary" (more generally finitely many) values. The probability $\{a_j = v_i\}$ before encoding the attribute value be

$$\text{Prob}(a_j = v_i|R) = P_i.$$

  It should be noted that, in the language of information theory, the binary credentials R correspond to the "information vector".
- The binary credentials corresponding to "age" attribute enable arriving at, say 1×4 information vector (corresponding to 16 age values). Thus

$$P_i = \text{Prob}(a_j = v_i|R) \quad (1)$$

  (repeated for the sake of clarity) is the conditional probability of $\{a_j = v_i\}$ under revealed credential set R (In this Example they are 16 binary vectors).
- Now the information vector (based on revealed credential set) is encoded into a codeword vector using a linear/non-linear error correcting code. For the sake of concreteness, let the codeword vector be a 1×7 vector (i.e. say a (7,4) Hamming code is used). This coding is done by using 3 credentials (say parity credentials) which assume only binary (more generally finitely many) values.
- It should be noted that the "age" attribute is encoded using binary credentials.

Case 1: The user who presents the "binary credentials" to estimate the "age" attribute will not be knowing the complete "error correcting code" utilized at the source Or equivalently.

Case 2: The encoding may be only known completely to the individual who presents the credentials and the system which tries to estimate the attribute does not know the encoding.

For the sake of concreteness, we consider Case 1 in the following discussion. It should be kept in mind that the following discussion also applied to Case 2.

Thus, let the probability of $a_j = v_i$ by presenting the credentials $C^*_{i_1}, C^*_{i_2}, \cdots, C^*_{i_k}$ be

$$P^*_i = \text{Prob}(a_j = v_i|R \cup C^*_{i_1} \cup C^*_{i_2} \cup \cdots \uplus C^*_{i_k}).$$

Now with the above description, we have

$$H(X) = \sum_{i=1}^{k} -P_i \log_2 P_i$$

entropy of source modelling the "AGE" attribute.

- Let Y be the random variable modelling the credential vector presented by the user (who does not know the encoding used by the transmitter). Thus

$$H(X|Y) = \sum_{i=1}^{k} -P^*_i \log_2 P^*_i.$$

- The process of presenting a set of credentials by a user to estimate the "age" attribute corresponds to transmitting the information vector through a Discrete Memoryless Channel. The stochastic matrix describing the channel (i.e. the channel matrix) is given by

$$Prob(X = v_i|Y = v_j)P_{i,j} \text{ for } i, j = 1, 2, \cdots, k. \quad (2)$$

Goal of The Modelling Problem:

Using the modelling assumptions described above and using the modified Equation (1) in Zhong and Bhargava's paper [17], we have Information in X provided by

$$\begin{aligned} Y &= \sum_{i=1}^{k} -P_i \log_2 P_i - \sum_{i=1}^{k} -P^*_i \log_2 P^*_i \\ &= H(X) - H(X|Y) = l(X;Y). \end{aligned} \quad (3)$$

i.e. $l(X;Y)$ denotes the mutual information between random variables X and Y.

Thus, so far **we have formulated the privacy-trust tradeoff problem as the problem of transmitting the attribute through a noisy communication channel (treated using information theory. It should be kept in mind that the process of guessing the attribute value using credentials is equivalent to the problem of transmission of information vector through a noisy channel).**

The above formalization/translation enables us to invoke the results from information theory. Precisely, we infer that as long as "data rate (a measure related to privacy loss)" is below a certain limit (determined by the channel capacity of discrete memoryless channel), it is possible to achieve arbitrarily high TRUST/RELIABILITY in estimating the attribute value.

By definition:

$$\text{C} = \text{Channel Capacity} = Maximum_p I_p(X;Y).$$

i.e. Maximum value of mutual information over all input probability distributions. Now we invoke the Shannon's channel coding theorem.

Theorem 1 *Given a discrete memoryless channel with capacity $C > 0$ and a positive number $R < C$, there exist a sequence of block codes $A_1, A_2, \cdots, A_n$ such that (on utilization of those codes for transmission) the maximum probability of error is as low as desired.*

Thus it is possible, by choosing n sufficiently large, to reduce the maximum probability of error to a figure as low as desired while at the same time maintaining the transmission rate R.

With the proper translation, the above theorem provides fundamental limitations on privacy-trust tradeoff problem for the case of query independent privacy loss.

Now we consider the next problem.

- **Query-Dependent Privacy Loss:**
  In this case, based on query 'q_k', the domain of an attribute "a_j " is divided into 'r' subsets i.e. $\{qv_1^j, qv_2^j, \cdots, qv_r^j\}$ based on the query answer set From Equations (2) and (3), for each query "q_k"; the privacy loss is measured as the difference between entropy values. Thus it indicates the mutual information indexed by query "q_k".

- Thus, the approach used in the case of query independent privacy loss also extends easily to this case. The discussion is avoided for brevity.

As in [17], using pr_i as the probability that query q_i is asked and w_i as the corresponding weight, the query-dependent privacy loss with respect to attribute "a_j" is evaluated.

In summary, since pr_i and w_i are independent of the problem of trading privacy with respect to attribute "a_j"; the approach discussed previously also extends to this case.

Generalization: (Invoking Multi-User Information Theory)

With the above results; we are naturally led to the **problem of trading privacy of multiple users**. Once again, using the results from multi-user information theory; basic limitations on privacy-trust tradeoff are easily derived.

5 Practical Implementation Issues

- As mentioned in [17], the determination of conditional probabilities is application specific. The techniques like Bayes networks [8] and kernel density estimation can be considered for the estimation problem.

- A large body of research literature exists on the design of linear/non-linear error correcting codes which are used to detect and correct multiple errors arising in transmission of an information vector. Those results are readily extended to the problem of trading privacy for "trust".

- If the error correcting code utilized for encoding the attribute value/subset is completely revealed to the receiver (to show the perfect "credentials"), the attribute can be PERFECTLY estimated. Thus absolute trust is established.

Noise occurs due to the following reasons:

1) The binary credential vector/set shown by the receiver does not match the one used by the transmitter. Thus the coding utilized by the source is not known (By adding arbitrarily large amount of redundancy i.e. parity bits, privacy is compromised).

2) The attribute value is coded using a certain error correcting code and the receiver only knows part of it.

3) The receiver does not know the error correcting code used at all (complete ignorance).

The results discussed can easily be extended to the case where the credentials are encoded using a non-binary (possibly non-linear) error correcting code.

6 Conclusions

Several authors [1, 13, 17] attempted to quantify concepts like privacy, anonymity etc. using some metrics. Using entropy related privacy metrics and the idea of issuing credentials to compromise privacy, the tradeoff is interpreted using the results of information theory. Since the modelling assumptions are very realistic, it is hoped that the results in the paper are of practical utility in designing schemes to trade privacy for trust in open environments like Internet.

Acknowledgements

The author would like to thank Dr. Yuhui Zhong and Dr. Leszek Lilien for providing the reference [17].

References

[1] D. Agrawal and C. Agrawal, "On the design and quantification of privacy preserving data mining algorithms," *Symposium on Principles of Database Systems (PODS'01)*, pp. 247-255, Santa Barbara, May 2001,

[2] S. Ba, A. B. Whinston, and H. Zhang, "Building Trust in the electronic market through an economic incentive mechanism, "in *Proceedings of 20th International Conference on Information Systems*, pp. 208-213, Dec. 1999.

[3] B. Bhargava and Y. Zhong, "Authorization based on evidence and trust," in *Proceedings of Intl. Conference on Data Warehousing and Knowledge Discovery (DaWak)*, France, Sept. 2002.

[4] J. Carbo, J. M. Molina, and J. Davila, "Trust management through fuzzy reputation," vol. 12, pp. 135-155, World Scientific, Mar. 2003.

[5] S. Farrell and R. Housley, *An internet attribute certificate profile for Authorization*, http://www.ietf.org/internet-drafts/draft-ietf-pkix-ac509prof-09.txt

[6] K. Fujimura and T. Nishihara, "Reputation rating system based on past behavior of evaluators," in *Proceedings of the 4th ACM Conference on Electronic Commerce*, pp. 246-247, ACM Press, 2003.

[7] I. Goldberg, *A pseudonymous communications infrastructure for the Internet*, Ph.D. thesis, University of California at Berkeley, http://www.isaac.cs.berkeley.edu/ iang/thesis-final.pdf, 2000

[8] F. V. Jensen, *An introduction to Bayesian networks*, UCL Press, London 1996.

[9] M. Reiter and A. Rubin, "Crowds: Anonymity for web transactions," *Communications of the ACM*, vol. 42, no. 3, pp, 32-48, 1999.

[10] P. Resnick, K. Kuwabara, R. Zeckhauser and E. Friedman, "Reputation systems," *Communications of the ACM*, vol. 43, no. 12, pp. 45-48, 2000.

[11] C. K. Sandberg, "Privacy and customer relationship management: Can they peacefully co-exist," *William Mitchell Law Review*, vol. 28, no. 3, pp. 1147-1162, 2002.

[12] J. M. Seigneur and C. D. Jensen, "Trading privacy for trust," in *iTrust'04 the International Conference on Trust Management*, LNCS, Springer-Verlag, 2004.

[13] A. Serjantove and D. Danezis, "Towards an information theoretic metric for anonymity," in *Proceedings of Workshop on Privacy Enhancing Technologies (PET'02)*, LNCS 2482, Springer-Verlag, 2002.

[14] L.Sweeney, *Computational disclosure control: A primer on data Pprivacy protection*, Ph.D Thesis, MIT, 2001.

[15] W. Wagealla, M. Carbone, C. English, S. Terzis and P. Nixon, "A Formal Model of Trust Lifecycle Management," in *Workshop on Formal Aspects of Security and Trust (FAST2003)*, 2003.

[16] T. Yu, M. Winslett, and K. E. Seamons, "Supporting structured credentials and sensitive policies through interoperable strategies for automated trust negotiation," *ACM transactions on Inf. Syst*, Secur, vol. 6, no. 1, pp. 1-42, 2003.

[17] Y. Zhong and B. Bhargava, "Using entropy to trade privacy for Trust," in *Proceedings of Security and Knowledge Mangement (SKM), 2004*, Amherst, NY, Sept. 2004.

[18] Y. Zhong, Y. Lu, and B. Bhargava, *Tera: An authorization framework based on uncertain evidence and dynamic trust*, Tehcnical Report, CSD-TR 04-009, Department of Computer Sciences, Purdue University, Feb. 2004.

Garimella Rama Murthy received Ph.D. degree in computer engineering from Purdue University, West Lafayette. He then worked as a member of technical staff at Bell Communications Research. He then consulted for Qualcomm and other companies in USA. He has been an associate professor at the International Institute of Information Technology (IIIT), Hyderabad for the past four years. He was a member of Phi Kappa Phi, Eta Kappa Nu etc. He is currently a member of IEEE and ACM. Dr. Rama Murthy received First Prize in All India Technical Essay Competition organized by the Visvesvaraya Industrial and Technological Museum, Bangalore. He also received 3rd Prize in All India Electronics Design Competition organized by the Institution of Engineers, India. His research interests are in Novel Co-ordinated Machine Architectures, Wireless Networks, Neural Networks, etc.