Malafide Intension and its mapping to Privacy Policy Purposes for Masquerading

Vikram Goyal
Comp. Sci. & Engg. Dept.
I.I.T. Delhi
Hauz Khas, New Delhi-16
vkgoyal@cse.iitd.ac.in

Shyam K. Gupta
Comp. Sci. & Engg. Dept.
I.I.T. Delhi
Hauz Khas, New Delhi-16
skg@cse.iitd.ac.in

Anand Gupta
Comp. Sci. & Engg. Dept.
N.S.I.T. Delhi
Sector-3, Dwarka, New Delhi
anand@coe.nsit.ac.in

Abstract

In presence of a robust privacy infrastructure, an attacker can fulfill his purpose(malafide intension) only by masquerading it with bonafide purposes besides other authentication parameters. We address the issue of masquerading of purpose for a malafide intension by defining the mapping from a malafide intension to bonafide purposes in this paper. An understanding of such a mapping can facilitate both a hacker (assist him in masquerading) and a forensic expert to investigate malafide accesses. Determination of these bonafide purposes may help speed up the violation detection if the User Accesses Log has listed bonafide purpose with each user access. The bonafide purposes can be determined in data-independent (without accessing the database) or data-dependent (database access is required) mode. In this paper we define a mapping of a malafide intension to bonafide purposes in data-independent mode.

1. Introduction

Privacy is of paramount importance in e-commerce, banking and plethora of web services. Many authors have proposed privacy policies[3, 1]. In-spite of lot of work being done towards providing robust privacy[5, 2, 6], there has been a steady rise in breach of privacy[4, 9]. This necessitates the need of post-event detection system[8, 7] which takes into account this factor. We believe that one reason for the failure of privacy frameworks is masquerading of authentication parameters like bonafide purpose by the intruder.

We define the specified purpose of access in privacy policies as *intension*. An intension is termed *bonafide* if the purpose of the user access is same as the purpose defined for the access by the information system. If it doesn't then it is termed *malafide*. In presence of a robust privacy infrastructure an access can only succeed through a bonafide purpose. A malicious user with a specific malafide purpose therefore would have to masquerade his malafide purpose to a set of bonafide purposes to gain access to the system. All the users who access the same information which has been accessed by a privacy violator with some malafide intension are termed as *malafide users* (inclusive of privacy violator) and their accesses as *malafide accesses*.

We hypothesize that every privacy intrusion has some malafide intension associated with it and that intension may be available after the violation has occurred. It can be associated with intrusion characteristics i.e. target data of the intruder. We claim that if malafide intension is available after the violation it can play a positive role in identification of malafide accesses and malafide users efficiently.

2. Malafide Intension

We define the semantics of malafide intension as the *masqueraded purpose* of the intruder. It can be associated with the information of the privacy violation in the post event scenario e.g. target data, quantity of data, time duration etc.. Target data in relational database can be represented by a set of set of *SQL* queries i.e. . Use of these semantics will help to identify the masqueradable privacy policy purposes(bonafide) which can be used to access the target data. We present mapping of a malafide intension to a set of bonafide purposes in this paper.

3. Mappings of Malafide Intension to a set of Bonafide Purposes

Every data retrieved from an information system has a purpose hence it is imperative that privacy policies state the purposes for which data is used or is going to be used. These purposes mentioned in the privacy policies are called bonafide purposes. A bonafide purpose is associated with a set of *data-items* that can be used/accessed for the purpose. A data-item can be represented by a set of *SQL* queries.

10th International Database Engineering and Applications Symposium (IDEAS'06)
0-7695-2577-6/06 $20.00 © 2006 IEEE

The mapping between a malafide intension and bonafide purposes can be defined on the basis of overlapping of information between target data of malafide intention and data items of bonafide purposes. As both the data (target-data and data-item) have SQL representation, the data-independent mapping(analysis over SQL syntax) would give false positive. For example, let T_1 is the target data with a malafide intension m_1 where T_1 is "*select* name, ssn *from* Table1 *where* region='x' and salary $\geq$ 3000", and D_1 and D_2 are data items associated with a bonafide purpose p_1 and p_2 respectively where D_1 is "*select* name, ssn *from* Table1 *where* region='x' and sex='m' " and D_2 is "*select* name, ssn *from* Table1 *where* region='x' and salary $\geq$ 2000". Then it is trivial to see that D_2 contains T_1, but there may or may not be some common records between D_1 and T_1. To test whether there is some data common between D_1 and T_1, data dependent mapping is required in which queries are executed over the database and their results are checked for overlapping.

Besides data overlapping, the mapping function will also depend on the attack model e.g. (a) whether single bonafide purpose is used to make attack or (b) a set of bonafide purposes is used to make the attack, (c) whether the entire data specified in target-data is required or (d) partial data from the target data can fulfill the malafide intention.

We define mapping of a malafide intension to a set of bonafide purposes as a function $map : M \rightarrow 2^{2^{2^B}}$ which determines the set of purposes set through which a data-objects of a target data of the malafide intension can be accessed.

Let $B = b_1, b_2, \cdots, b_k$ is the set of bonafide purposes. Let f_{bd} be a function which maps a bonafide purpose to a data-item and f_d be a function which maps a malafide intension m_j to the target data. Target data for a malafide intension is a set of data-objects. We use the notation $f_{bd}(S) = \cup_{b_j \in S} f_{bd}(b_j)$, where S is a set of bonafide purposes.

We define mapping function for attack model *(a,c)* as
$map(m_i) = \{f_{map}(o) | o \in f_d(m_i)\}$ where
$f_{map}(o) = \{p | p \in P \wedge p \text{ covers } o\}$ where
$P = 2^{f_{ob}}$ and
$f_{ob} = f_d(m_i) \rightarrow 2^B : f_{ob}(o) = \{b_i | b_i \in B \wedge o \subset f_{bd}(b_i)\}$,
p covers o means $f_{bd}(p) \supset o$.

For attack model *(b,c)* the definition of f_{ob} will change to
$f_{ob}(o) = \{b_i | b_i \in B \wedge o \cap f_{bd}(b_i) \neq \phi\}$.

For attack model type *(b,d)* the definition of f_{bd} will be same as for attack type *(b,c)* but the definition of *covers* will change to "p covers o means $f_{bd}(p) \cap o \neq \phi$".

For the example figure 1, the mapping for (a,c) would be $\{\{\{b_1\}\}, \{\{b_1\}, \{b_3\}\}, \{\{b_4\}\}\}$, the mapping for (b,c) would be $\{\{\{b_1\}\}, \{\{b_2, b_3\}, \{b_1, b_2, b_3\}\}, \{\{b_4\}, \{b_3, b_4\}\}\}$, and mapping for (b,d) would be $\{\{\{b_1\}\}, \{\{2^{\{b_1,b_2,b_3\}} - \phi\}\}, \{\{2^{\{b_3,b_4\}} - \phi\}\}\}$.

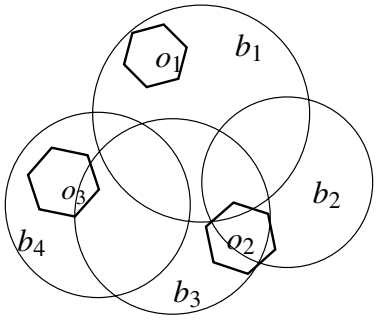

Figure 1. Mapping of Malafide Intension to Bonafide Purposes

4. Acknowledgment

We thank Department of Information Technology, Government. of India for funding "Design and Development of Malafide Intension based Privacy Violation Detection" project.

References

[1] IBM Tivoli Privacy Manager for e-business. http://www-306.ibm.com/software/info/ecatalog/en_TH/products/K106003J38182X80.html.

[2] OASIS, eXtensible Access Control Markup Language (XACML) Version 1.1. OASIS, 07 August 2003.

[3] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu. Hippocratic databases. In *Proceedings of the 28th VLDB Conference*, Hong Kong, China, 2002.

[4] J. B. Bruno. Security Breach Could Expose 40M to Fraud. Associated Press, June 18 2005.

[5] S. M. et al. Enterprise Privacy Authorization Language (EPAL 1.1). http://www.zurich.ibm.com/security/enterpriseprivacy/epal. IBM Research Report.

[6] V. Goyal, S. Gupta, I. Meshram, and A. Gupta. PRINDA: Design and Implementation of Non-Disclosure Agreement in Privacy Policy. April 3-7, 2006.

[7] V. Goyal, S. K. Gupta, S. Saxena, S. Chawala, and A. Gupta. Query Rewriting for Detection of Privacy Violation through Inferencing. In *International Conference on Privacy, Security and Trust, Oct30-Nov1, 2006*, 2006.

[8] S. K. Gupta, V. Goyal, and A. Gupta. Malafide Intension based Privacy Violation Detection. In *International Conference on Information Systems Security (ICISS 2006)*, Kolkata, India, 2006.

[9] B. Teasley. Does Your Privacy Policy Mean Anything? http://www.clickz.com/experts/crm/analyze_data/article.php, January 11 2005.