Improving Data Security and Privacy based on a Fuzzy Logic Classifier

Masoud Mohammadian
Faculty of Information Science and Engineering
University of Canberra
Canberra, ACT, Australia
masoud.mohammadian@canberra.edu.au

Dimitrios Hatzinakos
Electrical & Computer Engineering
University of Toronto
Toronto, Canada
dimitris@comm.utoronto.ca

Data security and privacy are very important issues in the success of a business operation. Implementing and applying policies related to data security and privacy therefore has become one of the core and important activities for large organizations.

Data classification process allows organizations to organize their data according to their needs.

This process can be laborious in large organizations with significant data to evaluate and categorize.

Using a data classification process organizations, can identify and apply appropriate policy and security settings such as private access control and encryption requirements.

This research study explores the use of fuzzy logic [1] in classification of data and suggests a method that can determine requirements for data security and privacy in organizations based on organizational needs and government policies imposed on data.

Sensitive and financial mission critical data are stored in databases, in server applications and/or middleware and data encryption at this level although useful can be disruptive and costly.

Managing the keys for encrypted data can become cumbersome and therefore many large organizations choose to encrypt only their regulated data [2, 3, 4, 5].

Using a data classification process financial organizations can identify and encrypt only the relevant data.

This will assist in saving time and processing power that is required for encryption and decryption process. [2, 3, 4, 5].

To classify data with minimal resources impact and without needing to re-design databases one option is to add extra information to each data item by adding meta-data information to the attributes of each entity in relational databases.

These meta-data information could be the value or degree of security, privacy or other related policies for that data item.

Case Study:

Consider the following entities of a relational database system:

Customer (CustomerID, Name, Address, TelNo, E-mail)

Product (ProductID, Name, Size, Color, Price)

Supplier (SupplierID, Name, Address, TelNo, FaxNo, E-mail)

Order (OrderID, CustometID, ProductID, SupplierID, OrderDate, Quantity)

Metadata value (related to security attributes for attributes) in tables above are based on organization's security policy and government security and privacy policy.

Assume that domain meta-data values for these linguistic variable are, **TP** = top secret, **SE** = "secret", **CO** ="confidential", **MC** = "mission critical", **NC** = "not critical", **PR** = "private but not top secret", **PU** = "Public". The values related to linguistic variables are: **TP** = [58,..,70], **SE** = [48,..,60], **CO** =[37,..,50], **MC** = [28,..,40], **NC** = [16,..,30], **PR** = [8,..,20], **PU** = [0,..,10].

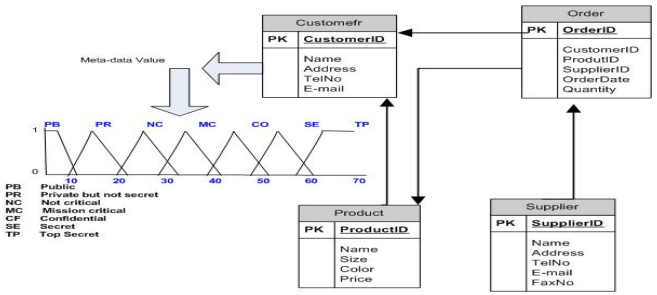


Figure 1. A relational database with metadata vales and its associate fuzzy set

Meta-data Value base on:

	Organization Policy	Government Regulatory Policy
CustomerID	**68**	**39**
Name	**64**	**70**
Address	**30**	**60**
TelNo	**44**	**68**
E-mail	**67**	**69**

Based on the metadata value for each attribute the membership of that attribute to each linguistic variable can be calculated.

Fuzzy membership of metadata value of CustomerID based on:

(a) Organization policy

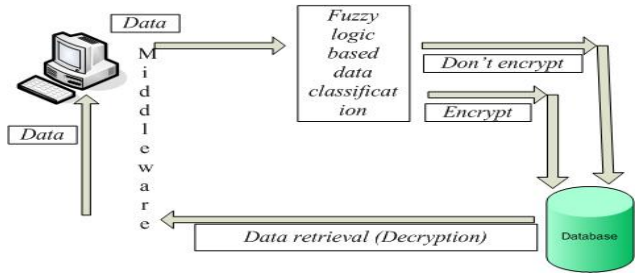
	TP	SE	CO	MC	NC	PR	PU
μ(CustomerID)	**0.8**	0	0	0	0	0	0

(b) government regulatory policy

	TP	SE	CO	MC	NC	PR	PU
μ(CustomerID)	0	0	**0.3**	**0.16**	0	0	0

Now that the data can be classified and categorized into fuzzy sets (with membership value), a process for determining precise actions to be applied must be developed using a fuzzy rule-based system.

IF *Organizational Security Classification* is TopSecret and *Government Security Classification* is Confidential Then *Level of Encryption required is* High



The proposed method in this paper provides a suitable data classification based on fuzzy logic for data security and data privacy.

Reference

1. L. A. Zadeh.,"Fuzzy sets", Information and control, Vol. 8. pp 338-352, 1965.

2. A. Spalka, J. Lehnhardt, "A comprehensive approach to anomaly detection in relational databases" DBSec, 2005, pp. 207–221.

3. H. H. Hosmer, "Using Fuzzy Logic to Represent Security Policies in the Multipolicy Paradigm", ACM SIGSAC Review, 1993.

4. S. De Capitani di Vimercati, S. Foresti, P. Samarati, "Recent Advances in Access Control", in Handbook of Database Security – Applications and Trends", M. Certz, S. Jajodia (Editors), Springer, USA, 2008.

5. Q. Yao, A. An, X. Huang, "Finding and analyzing database user sessions" Proceedings of the 10th International Conference on Database Systems for Advanced Applications (DASFAA), 2005.