

# Formalizing Authenticity and Security of Information in Information Processes

Navin Kumar Sehgal  
Department of Computer Science  
Kalindi College, University of Delhi  
New Delhi, India

Krishan Kumar  
Department of Computer Science  
Kalindi College, University of Delhi  
New Delhi, India

**Abstract**— The information systems are getting more and more complex causing a concern for the security of information and the processes involved. With advent of social networks and greater reliability on shared information, authentication has become a challenge. As information flows freely among the computing machines it becomes necessary to set standards for its creation, storage and usage. There is need for system that performs various information tasks while maintaining its security and authenticity.

**Keywords**—Information security, Information processes, Information authenticity.

## I. INTRODUCTION

According to empiricism our senses are the ultimate source of information. Various phenomenons around us help in creating information. In computer science sensors and related processes are a big source of information. Humans use rationalism, cognition and falsification techniques to process information. In this paper firstly we explore ways of information creation in secure and authentic manner. The method used is a feedback system which analysis threats with changing information patterns.

The next step is storage of information. It is a means of organizing information as well as storing it. Information can be classified according to content or value. Computer science relies on various methods for classification of information. In this paper we use Bayesian method of classification for categorizing information before storing it.

The final step involves understanding the usage of information. According to Taylor, 1990 information can be used for enlightenment problem understanding, instrumentation, confirmation, projection, motivation and personal. Multi level queuing model is used to study usage

## II. INFORMATION PROCESSES

Information has become a very useful resource in recent time. Formalization of information processes from creation to usage is a tool for authentication, security and optimization. The comparison can be made to digital camera which captures information, stores and makes further uses of it



Fig 1. Information flow across information processes

### A. Information Creation

This is the first step of information creation. Primitive information collected from sensors or added by a user. Information can be represented as documents or files.

The authenticity of information depends on the primitive information which is supposed to be recorded with unbiased sensors or humans. The information is secure if there is an access control mechanism to protect it. It can be calculated by ascertaining the degree of modifiability associated with created information.

If A and B are measures of security and authenticity Let P be the amount of information created into the system and Q is the information not obtained by shared resources. An increase in the value of P causes a change in values of A as well as B. Assuming k1 and k2 as constant for security measure and h1 and h2 as constants for authentication measure. This equation can be written as-

$$\frac{dA}{dt} = k1 * \frac{dP}{dt} + k2 * \frac{dQ}{dt} \quad (1)$$

$$\frac{dB}{dt} = h1 * \frac{dP}{dt} + h2 * \frac{dQ}{dt} \quad (2)$$

### B. Information Storage

Information created becomes meaningful if it is classified according to content and importance before storage. Manual classification involves grouping information according to given criteria. In a digital camera a user stores the images in various folders according to time or occasion or their projected frequency of use.

Bayesian method of classification is used for the problem. The baye's classifier consists of the probability model along with a decision rule. The method can be divided into training and classification phases. Information classification depends on the frequency of words used. Probability of occurrence of word is  $W_i$  is  $P(W_i)$ . Probability of word belonging to a particular class is  $P(W_i | X)$  and the probability that a word would belong to a class X is  $P(X | W_i)$ . It can be described in

Time (in days)	Total information (P Posts)	Shared information (Q Posts)	Creation step dA/dt	Creation step dB/dt	Storage step P(A)	Storage step P(B)	Usage step dA/dt	Usage step dB/dt
0	0	0	0	0	0	0	0	0
1	10	8	1.5625	1.5625	0.4080	0.5267	4.1667	1.3889
2	22	17	0.7353	0.7353	0.2395	0.3139	1.7544	0.6289
3	34	25	0.5000	0.5000	0.2102	0.2516	1.0309	0.4049
4	42	33	0.3788	0.3788	0.1115	0.1296	0.9524	0.3300
5	53	41	0.3049	0.3049	0.0987	0.1095	0.7299	0.2611
6	64	50	0.2500	0.2500	0.0763	0.0945	0.6173	0.2165
7	75	58	0.2155	0.2155	0.0699	0.0781	0.5155	0.1845
8	83	64	0.1953	0.1953	0.0644	0.0690	0.4630	0.1667
9	94	72	0.1736	0.1736	0.0594	0.0636	0.4032	0.1471
10	102	80	0.1563	0.1563	0.0465	0.0465	0.3906	0.1359

equation-

Table 1. Change in security and authentication values A, B with time (in hours)

$$P(X|Wi) = \frac{P(Wi|X) * P(X)}{P(X)} \quad (3)$$

### C. Information Usage

Information is used by people to fulfill various needs. According to Taylor information can be used for enlightenment or contextual purpose, problem understanding or comprehension of problem, instrumental or to do something, factual or gathering precise data, conformational or for verification purpose, projective or future oriented, motivational or work related, personal or relationship management.

The information is authentic if usage category is authentic. It can be calculated by relative authenticity of usage type. The information is secure if information is used securely. It is measured by amount of information used for secure usage.

A multi level queuing system is used for the purpose. Usage queues are assigned priority according to their authentic and security levels. If N1, N2, N3 and N4 are the four queues. The information entering each of these queues is from storage class C1 described as Q1c1, Q2c1, Q3c1, and Q4c1 and from storage class C2 as Q1c2, Q2c2, Q3c2 and Q4c2. Further N1 has the highest priority among all queues. Let A and B be the security and authentication parameters. Their values can be expressed in equation-

$$\frac{dA}{dt} = k1 * \frac{dQicj}{dt} \quad (4)$$

$$\frac{dB}{dt} = k1 * \sum \frac{dQicj}{dt} \quad (5)$$

### III. A CASE STUDY – SOCIAL NETWORKS

Before a study is carried out to evaluate information flow in software networks. A simulation based on the information processes is carried out. It is calculated separately for each of the processes

#### A. Understanding creation process

The study considers how information is created in a social network. The information is either an original post from a user or a shared post from another user. The number of shared and original posts increase with time. The security parameter measures the relative change in shared and original posts. The change is obtained after multiplying the changes in posts in two categories with k1 and k2. The values of k1 and k2 change for authentication measure.

#### B. Analysing the storage process

- Use Storage is based on classifying a post. Users usually categorize their posts into different lists. We have divided them into two classes as storage class C1 or storage class C2. Based on survey statistics values are assigned to C1 and C2. Bayesian classification is used to identify the probability that a post stored by the user in storage class C1 or storage class C2 is secure and authentic.

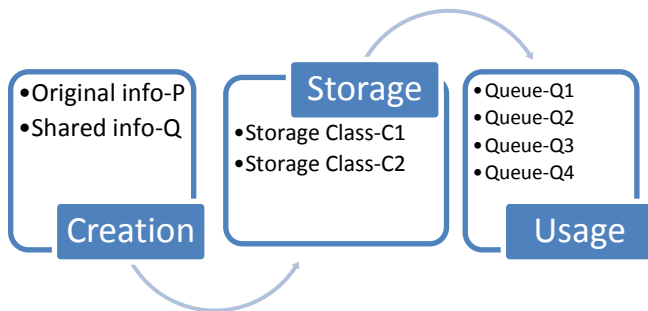


Fig 2. Information cycle for a social network

### C. Working of usage process.

The posts in social networks can be used for a number of purposes. Some users find the posts enlightening, others problem solving or helping them manage their relationships. We follow multi level queue scheduling the model. Assuming that information from the storage process enters any of the queues before usage, queues are categorized according to the purpose they serve. Confirmation and projection purpose usage is put in queue 1, Instrumentation and problem understanding purpose usage in queue 2, Enlightenment and motivational purpose usage in queue 3 and personal purpose usage in queue 4.

Assuming that queue 1 which stores posts to be used for projection and problem solving has the highest priority, it should be more secure as compared to other queues. The change in security parameter is calculated by measuring the relative change in secure and insecure information in queue 1. Authentication change is however calculated by measuring the overall change in information in all the queues.

### D. Benefits

- The study helps in analyzing behavior of social networks.
- It formalizes the press of movement of information from creation to usage.
- It creates a clear distinction between secure and unsecure information and their relative growth in a social network.
- It tries to analyze user behavior in using the information available on social networks and classifies it as authentic or non authentic.
- It uses Bayesian classification method to understand the probability of a post being unsecure and measures it for increasing volume of information.

## IV. CONCLUSION AND FUTURE WORK

As shown in table 1, we find that social networks become more unreliable as number of posts increases. There is a gradual decline in the security and authentication values in information creation process. Classification process also finds a large number of posts falling into the insecure class. The usage in turn becomes less authentic and secure due re sharing the same post over a large number of people.

The future work involves studying information processes to make them more secure. It can focus le on how to make social networks more reliable and authentic.

The methods used in this paper can be applied to a number of other applications like email system, document sharing system and location sharing applications

The future work can also focus on how a person should create and store information. It can benefit the areas where authenticity of information is a priority. Also it can benefit a person in helping him keep all his online information safe.

## REFERENCES

- [1] Nonaka, Ikujiro. "Managing the firm as an information creation process." *Advances in information processing in organizations* 4 (1991): 239-275.
- [2] Ugwuanyi, Richard NC, Michael C. Nwafor, and Chioma C. Onoh. "Networking of Academic Libraries: A Veritable Instrument for Information Creation and Dissemination."
- [3] Trace, Ciaran B. "Information creation and the notion of membership." *Journal of Documentation* 63.1 (2007): 142-164.
- [4] Rosenblatt, Frank. "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological review* 65.6 (1958): 386.
- [5] Korfhage, Robert R. "Information storage and retrieval." (2008).
- [6] Cutrell, Edward, and Zhiwei Guan. "What are you looking for?: an eye-tracking study of information usage in web search." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007.
- [7] Walters, Peter GP. "Export information sources—A study of their usage and utility." *International Marketing Review* 1.2 (1983): 34-43.
- [8] Allard, Suzie, Kenneth J. Levine, and Carol Tenopir. "Design engineers and technical professionals at work: Observing information usage in the workplace." *Journal of the American Society for Information Science and Technology* 60.3 (2009): 443-454.