I.J. Engineering and Manufacturing, **2017, 1, 1-9**
Published Online January 2017 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijem.2017.01.01

Available online at http://www.mecs-press.net/ijem

Relating Interactive System Design and Information Theory from Information Leakage Perspective

Kushal Anjaria, Arun Mishra

Defence Institute of Advanced Technology, Pune, India

Abstract

In contemporary interactive software system design, to maintain equilibrium between usability and security is a challenging task because strictly enforced security policies directly affect the usability of the software. As a solution to this problem, information theoretic measure of information leakage in interactive system design has been proposed in the present work. The present paper first models the software system as a coloured Petri net model and after that using information theory and Petri net algebra; it defines the leakage in the interacting system. Based on the leakage definition, the present paper further quantifies information leakage and tries to establish a relation between information leakage and interactive system design principles. The paper also hints to decide consensus on the equilibrium of security and usability.

Index Terms: Coloured Petri net, Information leakage, Information theory, Interactive system, Non-interference.

1. Introduction

The interactive system is a computer system where the user is prompted to provide input or information; in this way the interactive systems provides high usability to the user. But for security purpose, user provided input should not apprise him/her the secret information about the system. To achieve secrecy as well as usability, interactive system should at least satisfy following two conditions:

1. System should satisfy non-interference property[1,2]
2. System should follow interactive system design principles

But unfortunately both the conditions have limitations.
Consider the first condition. An interactive system satisfies non-interference if any sequence of low-security

* Corresponding author.
E-mail address:

sensitive user provided input will produce the same low-security sensitive output regardless of the high-security sensitive entity in the system. But it is very difficult to satisfy non-interference property for the interactive system because user provided information interacts or deal with the high-security sensitive entity of the system in one or the other way. Traditionally, non-interference has been studied as functional formulation. Therefore, the carefully built interactive systems also, do leak information. Consider the classical example of interactive login process which shows unavoidable leakage:

Example: The classical example of the interactive login process.
Scenario from attacker's perspective:

1. An attacker will interact with the system to gain access to the system.
2. The system will provide a dialog box to enter a password.
3. The attacker will guess the password as an input for login to gain access to the system.
4. Even if his guessed password is wrong, he will get this small information that his entered password is wrong.
5. Suppose initially attacker has k guesses for the password, after entering the wrong password, his guesses will be narrowed down to k-1.

The design principles mentioned in the second condition maintains the equilibrium between usability and security. But there is no consensus on how much strictly or leniently the principles need to be followed. In the present paper, two principles, principle of data display and constraint on user interaction have been considered.

In [13] and [14] the concept of interactive system design verification and validation has been discussed. For system designers to check that interactive system being designed is actually corresponding to requirement elicited during initial phase can be decided using a formal model of the user task. Most of the times, formal modelling of user action notation is done using Petri net model [14]. The work presented in the paper hints that the secrecy of the interactive system design can also be checked using colour Petri net [6] at this stage, which can add more value to the verification and validation process of the design of the interactive system.

1.1 Contributions:

The present work tries to solve the problem mentioned in above two conditions using quantitative analysis of leakage using coloured Petri net (CPN) approach. The compositional reasoning and Petri net based concurrency algebraic principles are the main motivating factors to use CPN based theories for the work presented in the paper. The use of Petri net in verification and validation stage of the interactive system design also motivates us to use CPN so that more gain can be achieved at verification and validation stage. The main contributions of the paper are:

- It models the interactive system as CPN and provides the definition of leakage in an interactive system using information theory.
- It tries to establish a relation between interactive system design principles and information theory. By using this relation, the paper provides a hint for determining the threshold of leakage while applying system design principles.
- The paper also provides leakage probability by establishing a relation between system observables and attacker's observational power and guessing ability.
- Some of the theoretical concepts [9, 10] available in process algebra are made applicable in practice for interactive system design.

1.2 Organization of paper

The second section of the paper incorporates basic definitions and preliminary concepts for the better understanding of the paper. Section-3 relates interactive system design and information theory using coloured Petri nets and process algebra. Section-4 provides background and related work in the area of information theory and interactive systems. Section-5 of the paper incorporates conclusion and possible future work.

2. Preliminaries

In this section, some of the concepts which are required for better understanding of the paper are briefly described. In the present work, information theory is used to quantify information leakage. In general information theory is a scientific discipline which studies about transmission, processing, quantification, utilization and extraction of information. In information theory, information is measured using Shannon's entropy. Formal definition of Shannon's entropy discussed in [20] is as below:

Shannon's entropy: Shannon defined the entropy H of a discrete random variable X with values $\{x_1, \ldots, x_n\}$ and probability mass function P(X) as:

$$H(X) = E[I(X)] = E\,[-\ln(P(X))] \tag{1}$$

Here E is expected value and I is information content of X. Then entropy can be defined as

$$H(X) = \sum_{i=1}^{n} P(xi)I(xi) \tag{2}$$

To relate the information theory and interactive system design principles, in the present paper, colour variation of Petri net is used. Formal definition of Petri net [21] is as below:

Petri net: (i) N = (S, T; F) is called a Petri net iff

(a) S and T are disjoint sets (S-elements and T-elements, respectively),
(b) F ⊆ (S × T) U (T × S), F is called the flow relation,
(c) $\nexists t \in T\ \exists s \in S\ tFs \wedge sFt$

3. Relating Interactive System and Information Theory

In the present paper, Jensen's [5] representation of coloured Petri net has been used; so instead of colour, tokens are represented with types e.g. X types of token. The CPN model of an instance of interactive system has been shown in the figure below:

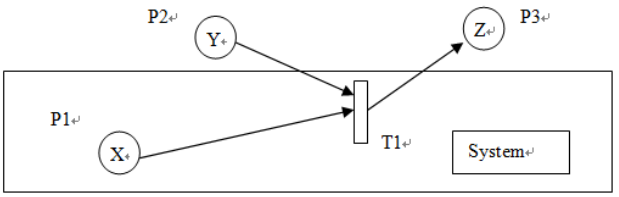

Fig.1. CPN Model of Interactive System

As shown in the figure, the system has an internal variable from variable set X = {xn| x is internal variable, n ≥ 0} shown by place P1. These variables will be represented with multi-set of x types of tokens. The variables in set X are high-security sensitive variables. Now onwards when variable x ϵ X is considered, automatically its corresponding token x from multi-set will be considered. Assume that system S satisfies the non-interference property.

The user interacts with the system and provides variable from variable set Y = {y_n| y is user provided variable, n ≥ 0} as his input shown as place P2. These variables will be represented with multi-set of y types of tokens. The variables in Y are low-security sensitive variables. The variables x and y may interact, transition T1 will fire and the system will produce output variable from variable set Z = {z_n| z is output variable, n ≥ 0} which is visible to the user. These variables will be represented with multi-set of z types of tokens. The variables in Z are low-security sensitive variables. Due to multi-set of tokens, places P1, P2 and P3 allow repetition of tokens which represent variables chosen from sets X, Y and Z respectively. For example, at place P2 user can enter token which represents variable y_1 ϵ Y multiple times in multiple instances. The variable sets are mapped with its random variable vectors respectively: X-P, Y-Q, and Z-R. Here vector P will hold all random variables of each element of set X. For example, random variable p_1 ϵ P will be associated with values that x1 can possess. Same applies for mapping Y-Q and Z-R. So there is one to one mapping between the elements of sets X, Y, Z and P, Q and R respectively.

At this point, the question arises that why the user provided input has been considered as a low-security sensitive entity. Closer observation of non-interference property reveals that actually partition of low and high-security sensitive variable is for formal use of the non-interference property. But in reality, property simply suggests that the outcome of the system should be independent of the input provided by the user and outcome of the system will be visible to all even to the attacker. So it should not reveal any secret. That's why in this paper users' input has been considered low-security sensitive which suggests that security level of output is same as input but system's internal entity is secret. This consideration is completely in sync with the non-interference property.

For the system shown in figure-1, places P1 and P2 can be considered as a creator of the message, the arcs can be considered as channels, the tokens can be considered as a messages and place P3 can be considered as a receiver. In this context, the interacting system can be considered as a communication system and can be studied from information theoretical perspectives. The transition T1 is a special transition whose firing rule is the presence of one token in either of the place P1 or P2 or at both the places P1 and P2. If the vectors of random variables are P, Q, and R then the definition of leakage function L(x,y), x ϵ X and y ϵ Y of system S can be given as below.

Leakage function: In an interactive system S, represented with CPN, if y type of token interacts with the x type of token and produce z type of token then leakage in interactive system S, due to user provided input interacting with internal variable can be described as a leakage function

$$L(x,y) = I(P; R|Q) = H(P|Q) - H(P|Q,R) \quad (3)$$

Where I describe mutual information, H is entropy [3].

Mutual information concept can be very clear from the Venn diagram shown below for three random variables. In [4] T. M. Cover and J. A. Thomas have described this diagrammatic concept of conditional mutual information.

From figure-2 the leakage definition becomes clear and the result I (P; R|Q) = H (P|Q) – H(P|Q,R) can also be derived. From the definition of leakage function and conditional mutual information, proposition is derived:

Proposition-1: In interactive system S represented with CPN if only x type of tokens are present at place P1 and token y at place P2 is absent then leakage function of resultant token z will be L(x,y) = H(R).

Proof:

$L(x, y) = I(P; R|Q)$ From the Venn diagram shown in figure-2 $I(P; R|Q)$ from R's perspective can be written as

$I(P; R|Q) = H(R|Q) - H(R|P,Q)$ But since Y is empty, vector Q is also empty. So, the entropy function $H(R|P,Q)$ will be 0 and $H(R|Q) = H(R)$.

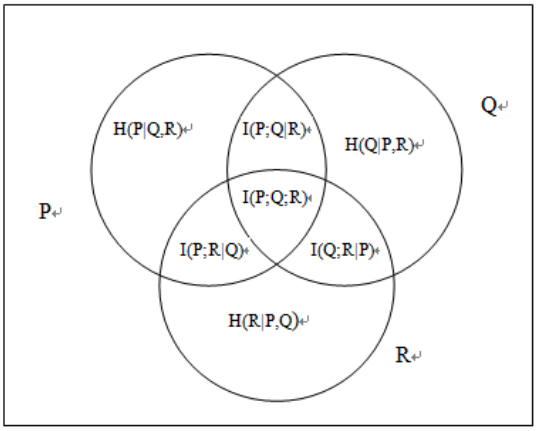

Fig.2. Venn diagram of Conditional Mutual Information

Note on proposition-1:

In the ideal interactive system, when Y is zero then leakage function $L(x,y) = H(R)$ should be nearly zero. That means if the user provided input is nil then the output should not leak any information about high security sensitive internal variable.

The proposition-1 supports one of the system design principles provided by Ben S. and C. Plaisant[7]: "Display the data only which assist the user, don't display more than that." The quantitative method presented in proposition-1 actually proves this principle. It shows that if this principle is violated then there is a chance that $L(x,y) = H(R) >> 0$. In other words, if the user is not interacting with the system (if Y=0) then system designer should carefully display the outcome of the system (i.e. Z). If in the outcome, more data is displayed, then using quantification method shown in proposition-1, it can be rectified. The proposition is in sync with the concept of absolute leakage in process calculus provided in [10]

3.1. Leakage analysis using Petri net equivalences

Behaviour equivalence of process using Petri net [8] can be defined as below:

Definition: 1 Let N_0 and N_1 are a two Petri net (P is set of place, T is a set of transition, M_{in} is initial markings, M is marking other than initial marking, $M_1 [t_1>M_2$ describes the transition from M_1 to M_2). A relation $R \subseteq$ Pow (P0) × Pow (P1) is behavioural equivalence between N_0 and N_1 (Pow is power set) if:

1. $(M_{in})_0$ R $(M_{in})_1$
2. if M_0 R M_1 and $M_0 [t_0>M'_0$, then there are $t_1 \epsilon T_1$ and $M'_1 \subseteq P_1$ such that $l_0(t_0) = l_1(t_1)$, $M_1 [t_1>M'_1$ and M'_0 R M'_1
3. as 2 but the role of N_0 and N_1 reversed
4. M_0 R $M_1 \rightarrow (M_0 = \Phi \leftrightarrow M_1 = \Phi)$.

From the observable equivalence, definition of secrecy [1] of interactive system can be given as below:

Secrecy of interactive system: An interactive system S(X,Y) keeps X secret if the outcome of the system doesn't depend on the different values that Y takes on.

System S shown in figure-1 represents only one instance of the interactive system. The user/attacker can repeatedly interact with the system. For example, in interactive login process if entered password is wrong, then the user can repeatedly enter the password. This will create instances. The outcomes of all instances will be behavioural equivalent. It will be behavioural equivalent because each instance will go through the similar sequence and outcomes of the sequences will have similar behaviour. The observable behaviour of the outcome of an interactive system can be formalized as behavioural equivalence as described in [12] using process algebra. But as shown in figure-1, the interaction in a system, shown as a communication system is among different-different entities represented as a node. And that's why the net representation of behaviour equivalence, described in definition-1 is better suited in this scenario.

Now, consider that there are n number of instances; user interacts with a system n times. These instances can also be represented with Petri net per instance. Consider an instance of system S shown in the figure-1, then a semantic [[.]] assigning to any instance S, meaning [[S]] is compositional and adequate if semantic equality implies congruence and behaviour equivalence for operator [9]. The instances of system S represented with the Petri nets will be behavioural equivalent as described in definition-1 because the system S satisfies non-interference property, the output Z of all instances will have same observable behaviour and it will not be dependent on Y as described in the definition of 'security of interactive system'.

Suppose instances of system S are represented as s_1, s_2, …, s_n.

Proposition-2 Let, semantic [[.]] has been assigned to all of its instances of system S. The Petri net of all instances are behavioural equivalent and S(X,Y) = [[$s_1(x_1,y_1)$, $s_2(x_2,y_2)$, …, s_n (x_n, y_n)]] then leakage of this semantic can be describes as $L(x_1,y_1) + L(x_2,y_2)+\ldots+L(x_n, y_n)$ and this sum of leakage will be greater than or equal to the leakage of entire system S

Proof:

As per the definition, leakage function is the function of entropy H.

In information theory, there is independence-bound theorem[4], which describes:

$$H(X1, X2, \ldots, Xn) \leq \sum_{i=1}^{n} H\ (Xi)$$

Same independence-bound theorem can be applied to leakage function, it is proved that $L(X,Y) \leq L(x_1,y_1) + L(x_2,y_2)+\ldots+L(x_n, y_n)$

Note on proposition-2

The proposition suggests that there should be a constraint on how many times a user can interact with the interactive system; more interaction generates more instances and leads to more leakage. For example, if the user can insert password without any constraints then the chances are high that user can guess the correct password. Therefore secure systems like ATM have constraints on the interaction of the user. This concept has been proved statistically in literature, but proposition-2 represents it in information theoretical way. The proposition-2 is supported by the principle of compositionality in process calculus provided in [10].

The proposition three relates attackers' observation power and equivalence.

Proposition-3 More discerning equivalence increases the power of the attacker to observe outcome and gain more about the secret from the outcome.

More observational power means an attacker can get more information about x from the observation of z. This result can also be proved easily using the definition of Petri net behavioural equivalence and equivalence

bound theorem. In the context of the system S presented in the paper, behavioural equivalence is natural because it is all about outcome z. In a practical system, the equivalence is the notion of the observer. So it is not about the choice of equivalence. In Petri net algebra theory, congruence, occurrence net equivalence, pomset equivalence, and trace equivalence are also present [8]. They can be used for various purposes like to find the execution path, chosen by the attacker, pomset or trace equivalence can be used. This proposition is a generalization of proposition-1. This proposition also suggests that user should not get more information about a secret from outcome and outcome should be independent of the user input. The quantification of the possible outcome is found using proposition-1.

The final proposition is about attacker's guessing of a secret in an interacting system.

Proposition-4: In an interactive system S(X,Y), the guessing of attacker about secret X from resulting Z can be described by guessing function

$$G(X) \geqq \frac{2H(X)+1}{2} \tag{4}$$

where G(X) is a guessing function and H(X) is an entropy function

The guessing function is introduced and proven by Massey [11] in his work. This type of guessing function will be useful in interactive system where an attacker can interact and query the system and in reply he get an answer in affirmation or negation, an example is interactive login system. Using this function, an interactive system designer can limit such interface where users get answers in affirmation or negation or when these types of interfaces are designed then the designer can measure the leakage. To the best of our knowledge, there is no design principle of interactive system stating user should be provided limited oracle. But the proposition-4 proves this and system designer may take this into the consideration for better security and avoid the oracle which gives an answer in affirmation or negation.

4. Related Work

For the first time, usability and system design were combined by J Gould and C. Lewis [26]. J Gould and C. Lewis have provided theoretical design principles for better usability; although, there wasn't any discussion on the consensus of usability, security and design principles. In [17], Sonia C. et al. have provided user pattern based approach for user interface design. The paper supports the graphical password based method for security. In [18], Marc A et al. have introduced UIML-User Interface Markup Language. The language facilitated security because it wasn't supporting procedural language. But while security facilitation, how it was affecting the usability and interactive system designing haven't been discussed. In [19] M. E. Zurko and R. T. Simon provided theoretical analysis to balance the usability, security and interactive system design. But the principles presented were theoretical and unable to design the crisp boundary between usability and security.

In [22], W. Newman has proposed new interactive graphical programming language. The proposed language was graphical problem oriented language. In [23], a framework for performance evaluation of interactive system was proposed by N. Beringer et al. The aim in [23] is to perform the evaluation of interactive system which has modal inputs and outputs. In [24] Donald B. discussed reliability issues in an interactive system in detail. In the work, Donald B. proposed resilient that capture the notion of security and reliability of a wide variety of interactive system. In [25], B. Gaines and P. Facey shared their own experience in developing interactive minicomputer systems in commercial, medical, industrial and scientific application. In [25] number of case histories and case studies are provided to share the experience of developing and designing interactive systems. In [26], web-based interactive system design and development were discussed by Zong-Ren Peng and Ruihong Huang. The interactive system proposed in [27] was web-based transit system which was using Geographic Information Systems (GIS) technologies to integrate Web serving, network analysis, and database management as an integral part of the interactive system. In [16], design issue of website sitemap feature has been surveyed and discussed. Web sites can also be considered and studied as an interactive system.

As per our knowledge, the present paper is the first attempt to mathematically provide consensus on the equilibrium of security, usability and system design principles.

5. Conclusion and Future Work

The paper presents the relation between interactive system design principle and information leakage. Based on the leakage quantification, the strict measure can be provided to the interactive system design principles. The present paper hasn't considered the cost of user interaction. In future, it can be assumed that user has limited chances to interact with the system. And for this purpose, traces of the user interaction and his success and failure chances can be quantified. The trace equivalence and testing equivalence will also be useful to find the rate of leakage of the interactive system. In this work, CPN is designed which handles alternate interactions of user and system. But in actual system interaction will not be alternate; for this type of system, CPN needs to be designed. In [15], a feedback channel is considered in information theory to quantify leakage in the non-alternate scenario of the interactive system. The present paper has represented two interactive system design principles to which quantitative measure are given, but there are other principles also which can be given quantitative threshold using the method proposed in the paper.

References

[1] Boudol, Gérard, and Ilaria Castellani. "Non-interference for concurrent programs and thread systems" Theoretical Computer Science 281.1 (2002): 109-130.

[2] Ryan, P., McLean, J., Millen, J., & Gligor, V. (2001, June). Non-interference: Who needs it? In csfw (p. 0237), IEEE.

[3] C. Shannon, "A mathematical theory of communication", The Bell System Technical Journal, volume 27, July and October, 1948, pages 379–423 and 623–656.

[4] T. M. Cover and J. A. Thomas, "Elements of Information Theory", 1991, Wiley Interscience.

[5] K. Jensen. Coloured Petri nets and the invariant method Mathematical Foundations on Computer Science, Lecture Notes in Computer Science, 118:327–338, 1981.

[6] Baracaldo, Nathalie, and James Joshi. "An adaptive risk management and access control framework to mitigate insider threats." Computers & Security39 (2013): 237-254.

[7] Shneiderman, Ben. Designing the user interface. Pearson Education India, 2003.

[8] Van Glabbeek, Rob, and Frits Vaandrager. "Petri net models for algebraic theories of concurrency." PARLE Parallel Architectures and Languages Europe. Springer Berlin Heidelberg, 1987.

[9] Jategaonkar, Lalita, and Albert Meyer. "Testing equivalence for Petri nets with action refinement: preliminary report." CONCUR'92. Springer Berlin Heidelberg, 1992.

[10] Boreale, Michele. "Quantifying information leakage in process calculi."Information and Computation 207.6 (2009): 699-725.

[11] Massey, James L. "Guessing and entropy." Information Theory, 1994. Proceedings., 1994 IEEE International Symposium on. IEEE, 1994.

[12] De Nicola, Rocco, and Matthew CB Hennessy. "Testing equivalences for processes." Theoretical computer science 34.1-2 (1984): 83-133.

[13] Mori, Giulio, Fabio Paternò, and Carmen Santoro. "CTTE: support for developing and analyzing task models for interactive system design."Software Engineering, IEEE Transactions on 28.8 (2002): 797-813.

[14] Palanque, Philippe, Rémi Bastide, and Valérie Sengès. "Validating interactive system design through the verification of formal task and system models." Engineering for Human-Computer Interaction. Springer US, 1996. 189-212.

[15] Alvim, Mário S., Miguel E. Andrés, and Catuscia Palamidessi. "Information flow in interactive

systems." CONCUR 2010-Concurrency Theory. Springer Berlin Heidelberg, 2010. 102-116.
[16] Manhas, Jatinder. "Comparative Study of Website Sitemap Feature as Design Issue in Various Websites." *IJEM-International Journal of Engineering and Manufacturing (IJEM)* 4.3 (2014): 22.
[17] Chiasson, Sonia, et al. "User interface design affects security: Patterns in click-based graphical passwords." International Journal of Information Security 8.6 (2009): 387-398.
[18] Abrams, Marc, et al. "UIML: an appliance-independent XML user interface language." Computer Networks 31.11 (1999): 1695-1708.
[19] Zurko, Mary Ellen, and Richard T. Simon. "User-centered security."Proceedings of the 1996 workshop on New security paradigms. ACM, 1996.
[20] Borda, Monica (2011). Fundamentals in Information Theory and Coding, Springer. p. 11. ISBN 978-3-642-20346-6.
[21] Goltz, Ursula, and Wolfgang Reisig. "The non-sequential behaviour of Petri nets" *Information and Control* 57.2 (1983): 125-147.
[22] Newman, William M. "A system for interactive graphical programming." *Proceedings of the April 30--May 2, 1968, spring joint computer conference*. ACM, 1968.
[23] Beringer, Nicole, et al. "Promise-a procedure for multimodal interactive system evaluation." *Multimodal Resources and Multimodal Systems Evaluation Workshop Program Saturday, June 1, 2002*. 2002.
[24] Beaver, Donald. "Foundations of secure interactive computing" *Annual International Cryptology Conference*. Springer Berlin Heidelberg, 1991.
[25] Gaines, Brian R., and Peter V. Facey. "Some experience in interactive system development and application." *Proceedings of the IEEE* 63.6 (1975): 894-911.
[26] Gould, John D., and Clayton Lewis. "Designing for usability: key principles and what designers think." Communications of the ACM 28.3 (1985): 300-311.
[27] Peng, Zhong-Ren, and Ruihong Huang. "Design and development of interactive trip planning for web-based transit information systems." *Transportation Research Part C: Emerging Technologies* 8.1 (2000).

Authors' Profiles

Kushal Anjaria is a Ph.D. scholar at the department of computer science and engineering of Defence Institute of Advanced Technology, Pune-India. He received M Tech in computer science and engineering from Manipal Institute of Technology, Manipal in 2012. His work is currently focused on information theory, information security, Petri net and quantitative analysis of information leakage. He can be reached at kushal_pcse14@diat.ac.in and kushal.anjaria@gmail.com.

Arun Mishra is an assistant professor at the department of computer science and engineering of Defence Institute of Advanced Technology, Pune-India. He got his Ph.D. in computer science from Motilal Nehru National Institute of Technology, Allahabad (India). His research activity is based on Automated Systems, Trusted Computing, Secure Software Engineering, Formal Modelling and Component based Software Engineering.

How to cite this paper: Kushal Anjaria, Arun Mishra,"Relating Interactive System Design and Information Theory from Information Leakage Perspective", International Journal of Engineering and Manufacturing(IJEM), Vol.7, No.1, pp.1-9, 2017.DOI: 10.5815/ijem.2017.01.01