Formalization and Quantitative Analysis of Controllability on Internet Information Security

Yin Lihua[1,4], Fang Binxing[1,2], Guo Yunchuan[1], He Hui[3]
[1]Institute of Computing Technology
Chinese Academy of Sciences
Beijing China
yinlihua@software.ict.ac.cn

[2]Beijing University of Posts and Telecommunications
Beijing China

[3]Computer Network and Information Security Research Center
Harbin Institute of Technology
Harbin China

[4]National Engineering Laboratory for Information Security Technologies
Beijing China

ABSTRACT: *Traditional information security models focus on protecting information or information systems. There is little research on controllability of Internet information security. We present network reference monitor (NRM) for access control in Internet environment. Then we propose the definition of network reference monitor and Internet information access control. Abstract operation model is put forward for controllability of Internet access control. The model consists of three parts: information discovery model, information analysis model and information control model. Next, we express the intuitive formal specification of controllability by PCTL logic. Then we characterize the process of Internet access control by finite state automata and present an evaluation model of controllability. Control Rate and Control Leak Rate are given to analyze access control of Internet information security quantitatively.*

Keywords: Information content security, Controllability, Quantitative analysis, Formal specification

Received: 12 September 2009, Revised 19 October 2009, Accepted 4 November 2009

1. Introduction

Network culture security is attracting more and more attention as Internet progresses. The principal part of network culture security is the information content security from the technological point of view [1]. Malicious information proliferates uncontrolled in Internet environment and does all kinds of harm. The shell viruses may ruin information receiving systems, the spam mails are a constant bother, the rumors puzzle the commonalty, and the eroticism harms the mental health of young people, etc. Thus the uncontrolled malicious information does endanger the public security seriously [2,3].

For forbidding promulgating pornography and illegal information in Internet, the UK promulgated the "R3 Safety Net", France the "Fillon Amendment", and Singapore the "Internet Code of Conduct". United States Congress passed "Communication Decency Act" and "Children's Online Privacy Protection Act" to forbid issuing malicious information. Nevertheless, the US Supreme Court ruled the two acts invalid for being unconstitutional. It was held that mental health of children should be protected by technological means rather than passing prohibitive laws against the issue of malicious information on the Internet [4, 5].

Traditional information security models protect information and/or information systems, since information is non-substantial and must be stored, or issued, or transferred, or processed in an information system. Some security properties of information, such as its confidentiality, integrity and availability, etc., are equivalent to security properties of information or information systems. Consequently, much attention is being paid to researches on security of information systems in academia.

However, information system security protects authenticity of information format and consistency of information transmission or storage. The research on information system security is aimed at protecting the information by protecting the grammar or format of information. But the information content security appraises the correctness or validity or authenticity of information by its content, that is, the meaning the promulgator actually wants to express. The research on information content security is aimed at protecting the information by protecting the semantic of information. From the technological point of view, information content security is the ability of selection and control with information flow. In other words, the controllability of information access process is emphasized in content security research [6].

As an important aspect of network security, access control is an increasing concern in academia. There have been developed corresponding access control models for different network environments that provide good network security solutions, and the researches were concentrated on concrete realization of access control for specific applications. But little work studies controllability as a security property on the theoretical level.

We study the reference monitor of access control in Internet environment for information content security. Then we present abstract operation models of controllability and formalize the property with intuitive formal specifications. On the basis of the models and formal specifications, we put forward a probabilistic evaluation model for controllability of content security access control.

The rest of this paper is organized as follows. Section 2 studies existing reference monitor and access control models and presents a network reference monitor for content security. Section 3 proposes an operational model of content security and intuitive formal definition and the evaluation method of controllability in Internet information access control. Finally section 4 presents conclusions and future work.

2. Network Reference Monitor

Internet information content security requests an ability to control the flow of information, e.g. certain information can be accessed and certain information can not. So Internet information content security can be modeled by access control. Access control is a mechanism to permit or deny the access of particular resources by particular entity. The resources, such as files, frames and packets etc, are referred to as objects collectively and the entities, which can access or use objects, are referred to as subjects collectively.

There have been different access control models for different applications. Those models can be divided into two categories, traditional access control models and modern access control models [7]. Traditional access control models include MAC, DAC and RBAC etc [8]. Modern access control models include $UCON_{ABC}$ [9], distributed access control [10], grid access control [11] and wireless network access control etc [12,13].

Reference monitor is responsible for the protection of accessed objects in access control model. Reference monitor stores access control rules and permits or denies the access process according to the rules. Communication initiator or information requestor sends request to ask for accessing objects which are network resources like files, equipment and/or CPU etc. Reference monitor checks the request and/or response and determines whether to allow the request and/or response to pass or not.

Reference monitors in existing access control are divided into two different categories [14]: Server Reference Monitor (SRM) and Client Reference Monitor (CRM) as shown in Fig.1 and Fig.2. SRM inspects whether the access process is authorized in server and CRM does the same in client.

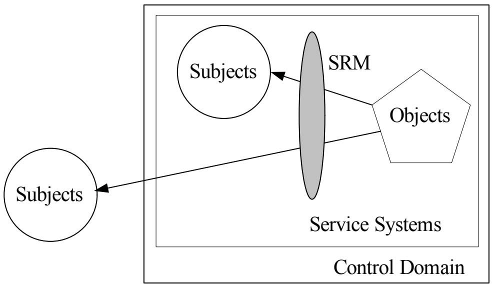

Figure 1. Server side reference monitor

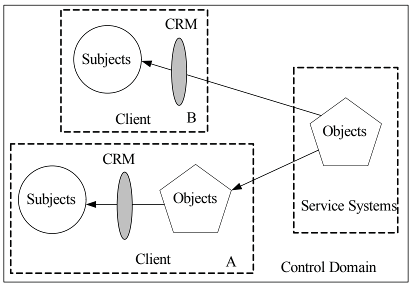

Figure 2. Client side reference monitor

However, in order to propagate harmful information, information propagators can not place SRM in servers to control accesses of information. Although some systems such as Internet Content Rating Association (ICRA) implement content security by CRM, users do not place CRM to prevent harmful information propagation in many cases and do not request RM to check the accesses.

We present Network Reference Monitor (NRM) to solve Internet information access control. According to the different placement, NRM has two types: gateway NRM and bypass NRM as shown in Fig.3 and Fig.4.

Gateway NRMs are placed on the critical path of network access, so they detect and transmit all network data package. The access requests are sent to NRMs in fact. The NRM analyzes the requests according to the control strategies and transmits the requests if they do not satisfy the rules. The responses are also analyzed by the NRM and the NRM determines whether to transmit the responses or send to the access requestor modified data or interrupt the access action.

Bypass NRMs are placed on the network main export in parallel to receive network traffic by mirror or beam splitter. The NRM analyzes the traffic data and sends control command to interrupt the access actions whose data matches control rules. The visitors will not feel the existence of NRM because bypass NRM is not located in transmission paths. And the normal traffic data between subjects and objects pass the network export directly.

Gateway NRMs have advantages of dealing with network access flexibly and having high success rate while filtering data. But massive information accesses in Internet need high-performance of gateway NRMs, and gateway NRMs easily become network transmission bottleneck or become the choice that attackers launch DoS attacks to cause a single point of failure. Bypass NRMs have advantages of not affecting the speed of network access and not bringing on single point of failure of network access. The shortcoming of bypass NRMs is inefficient on control of data content with sensitive information at the end of data stream due to the characteristics of bypass monitor.

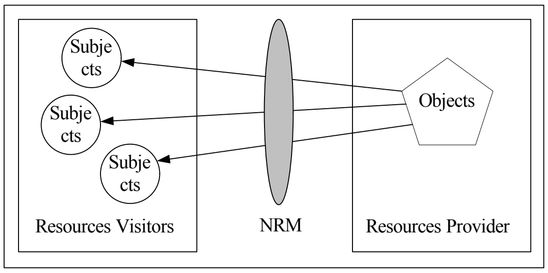

Figure 3. Gateway network reference monitor

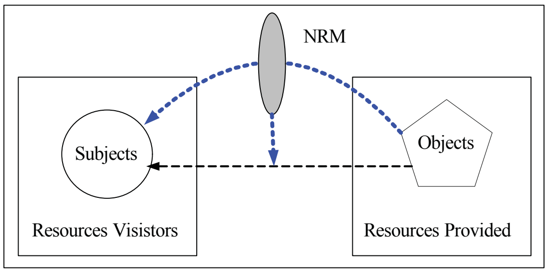

Figure 4. Bypass network reference monitor

3. Model and Formalization of Controllability

Controllability refers to ability to control users' access to network resources by access control strategies. For information content security, controllability refers to ability that authorized managers can preprocess and exert control on all access actions of web visitors.

3.1 Internet Information Access Control

Internet information access control mechanism controls access actions with static or dynamic control rules by NRMs.

Definition 1: Network Reference Monitor.

A network reference monitor is a 5-tuple:

$$\text{NRM} = < M, P, S, O, A >$$

where *M* is set of hardware and software, *P* is set of access control rules, *S* is set of access subjects, *O* is set of network resource objects , *A* is set of control commands.

The access control strategy *P* is constituted by authority and identity feature of subject, content feature of information and access action feature. The NRM controls actions of subjects *S* accessing objects *O* based on the access control rules *P* with control behaviors *A*.

Example 1: Consider a NRM, *M* denotes the hardware and software of a pornographic information control system, *S* consists of all students in campus, *O* consists of all pornographic information in Internet such as articles or pictures or videos, *P* consists of key pornographic information such as "oral sex" or erotic video segments or blacklist of pornographic URLs, *A* consists of control commands such as "ignore the request" or "discard the response page". The relationship of NRM's constituent is shown in Fig.5.

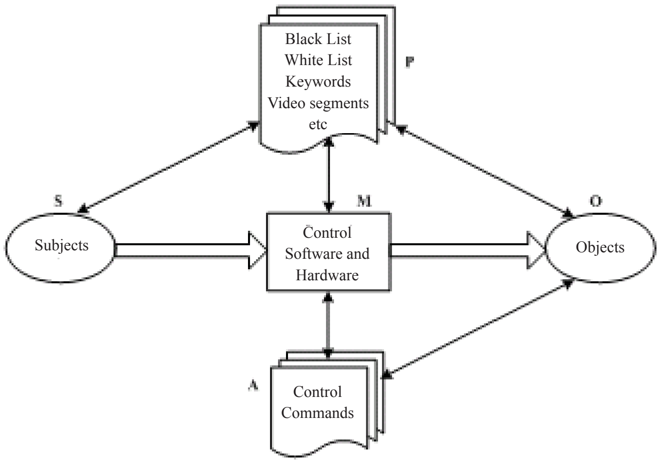

Figure 5. Relationship of constituent of network reference monitor

Definition 2: Network Information Access Control.

Whether an action a permitted or not is determined by satisfaction between control rules *P*(*X*) and network resource O which subject S wants to access. If resource O matches *P*(*X*), then the action a can not be permitted, else the action a is permitted.

Discriminant of control rules satisfiability is expressed by satisfaction (S⟨*f*(O)⟩, *P*(*X*)). The function value is true or false. *f*(*O*)denotes function *f* will be applied to network resources O, such as to read a page or download a file. Subject access action S⟨*f*(O)⟩ denotes subject S wants do some action f to network resources O, e.g. a user wants to upload some pictures to a FTP server. satisfaction (S⟨*f*(O)⟩, *P*(*X*)) denotes that subject access action S⟨*f*(O)⟩ satisfies control rules *P*(*X*). To the subject access actions S⟨*f*(O)⟩, the meaning of control is

$$\text{Control}\,(\text{S}\langle f\,(\text{O})\rangle, P(X)) \triangleq \text{if satisfaction}\,(\text{S}\langle f\,(\text{O})\rangle, P(X)) \text{ then } \neg f\,(\text{O}) \qquad (1)$$

Example 2: Consider a NRM with pornographic information control system, S consists of all students in campus, O consists of all pornographic information in Internet such as articles or pictures or videos, *P*(*X*) $\subseteq$ *P* is set of control rules consisting of key pornographic information such as "oral sex" or erotic video segments or blacklist of pornographic URLs. $f(O)$ denotes some actions will be done to objects O such as to download a file.

If there is an erotic video segment, which is contained in $P(X)$, in the file that user S wants to download, the download action $f(O)$ will not be permitted to implement. That is, if subject access actions $\mathrm{S}\langle f(\mathrm{O})\rangle$satisfies satisfaction $(\mathrm{S}\langle f(\mathrm{O})\rangle, P(X))$ and actions $f(O)$ is not implemented, we can say that access actions $\mathrm{S}\langle f(\mathrm{O})\rangle$ is controllable with $P(X)$, viz. Control $(\mathrm{S}\langle f(\mathrm{O})\rangle, P(X))$.

Information propagators and/or visitors want to escape monitor and/or control in order to propagate and/or access harmful information, so it is impossible that they request NRM to check the access information actively. At the same time, they may adopt anonymity technology and/or encryption technology to make uncertainty of correspondent identity and/or information content. Consequently, the access control of Internet information must involve how to get the information and discover the specific information.

3.2 Operational Model of Internet Content Security

The information issuance of Internet can be modeled as abstract information issuance model as shown in Fig.6. That is, information *ei* is injected into massive information aggregate *E* under condition *c*. viz. $f_{inject}(ei,c) \rightarrow E$, where information *ei* is original information issued, *E* is massive information aggregate of Internet, condition *c* is access condition while information is injected into massive information aggregate *E*.

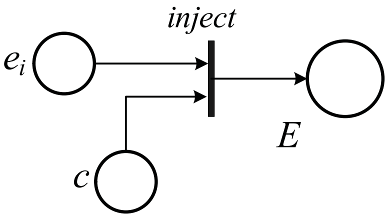

Figure 6. Information issuance operational model

The model is expressed by similar Petri net, Information *ei* and condition *c* and information aggregate *E* are symbolized by circles as the places of Petri Net and black rectangle is the transition of Petri Net. The transition f_{inject} is fired while the Place *ei* and *c* have resources and information *ei* is injected into aggregate *E*. Abstract operation models of information content security is presented after we abstract the inherent meaning of access control further.

Due to particularity of Internet information access control, Internet access control is divided into three phases: information discovery, information analysis and information control.

Information discovery is defined as to discover information *e* with specific content from Internet massive information aggregate *E*. There are two possible situations, one is information extraction with known extraction conditions $ej = f_{extract}(E,c)$ and the other is information mining with unknown extraction conditions $ej = f_{mine}(E,c)$. Information extraction can obtain original issued information and information mining may obtain information which has some certain characteristics.

Information analysis is defined as to determine the satisfiability of traffic data on control rules $P(X)$ with access control in NRMs. The determination method is abstracted as $f_{analyze}(e, P(X))$ and the determination result is true or false, that is, $f_{analyze} : E \times P \rightarrow \{0,1\}$. The abstract information content analysis operational model is shown in Fig.7.

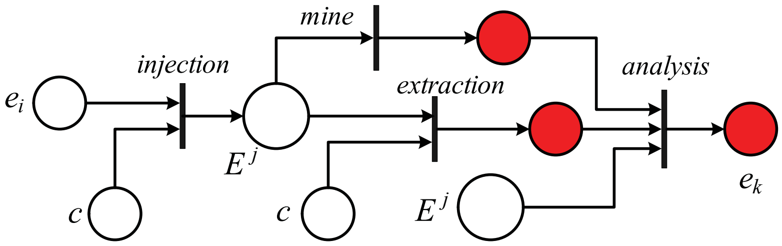

Figure 7. Information content discovery and analysis operational model

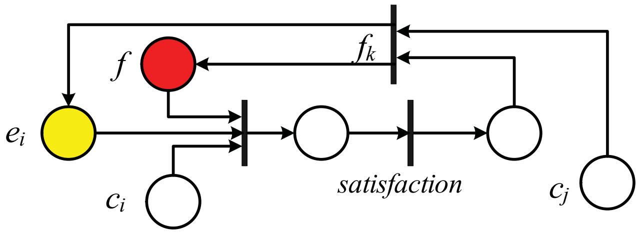

Figure 8. Information content control operational model

Information control is defined as function $f_{control}(x)$ controls original issued information or information access process according to determination result of satisfiability with control rules $P(X)$ under the condition c_j, where x refers to information e_i or access action a. The abstract information control operational model is shown in Fig.8.

3.3 Formal Specification of Controllability

There have been some researches related to formal specification of security properties for years. The typical researches can be classified into three types by description of specification: specification based on process algebras, specification based on epistemic logic and specification based on temporal logic.

Although process algebras have powerful express ability [15, 16], we can find that these methods are based on non-interference while they describe security properties. These methods are limited because non-interference is used to analyzing information leak. Epistemic logics are put forward to specify security properties in security protocols in order to analyze whether the system fulfills some specific properties [17]. Formal specification of security properties based on temporal logic formalizes authentication and confidentiality of security protocols or software requirements specification with CTL (Computation Tree Logic) or LTL (Linear Temporal Logic) [18, 19].

To support analysis and evaluation of network information access control, formal definition of controllability is needed. Therefore, we adopt PCTL (Probabilistic Computation Tree Logic), which is a logic supporting probability analysis, to formalize controllability. The grammar and formal semantics of PCTL refers to the literature [20].

According to the Internet information access control model, we suppose *extracted*, *mined*, *analyzed*, *matched*, *mismatched* and *controlled* as atomic proposition, which denote the process of extracting information from Internet, mining information from Internet, analyzing information, matching information with control rules, controlling specific information or access action correspondingly. Then controllability of information content security is defined as follow.

$$\begin{gathered}\Box(((extracted \vee mined) \wedge mismatched) \vee \\ ((extracted \vee mined) \wedge matched \rightarrow \Diamond controlled)))\end{gathered} \tag{2}$$

Where, $\Box$ denotes "always" and $\Diamond$ denotes "eventually". So, the formula above expresses that Internet information is controllable if information can always be extracted/mined from Internet and mismatch control rules or be extracted/mined from Internet and match control rules but be controlled eventually.

Controllability of Internet information content security is defined probabilistically as follow.

$$\begin{gathered}[\Box(((extracted \vee mined) \wedge mismatched) \vee \\ ((extracted \vee mined) \wedge matched \rightarrow \Diamond controlled)))]_{\geq p}\end{gathered} \tag{3}$$

That is, controllability is satisfied while probability of information, which always be extracted/mined from Internet and mismatch control rules or be extracted/mined from Internet and match control rules but be controlled eventually, is greater than or equal to p.

As formal description of controllability put forward above, controllability is ability of processing and controlling for actions which users access network resources according to access control rules. We use symbol S to denote set of network access subjects, symbol O to denote set of network resources and predication $C(S)$ to present controllability of access actions, and then the semantic of controllable network access action is expressed as follow.

$$\forall s \in S, \forall o \in O : \text{satisfaction}(s\langle f(o)\rangle, P(X)) \rightarrow \neg f(o) \models C(S) \tag{4}$$

where symbol $|=$ denotes the satisfaction relation which network access actions are permitted only when information content do not match the control rules *P(X)* for any action that subjects *s* access resources *o*.

Controllability measurement is to evaluate probability of controllable network access actions. So controllability measurement can be expressed by formula as follow.

$$\forall s \in S, \forall o \in O : \text{satisfaction}(s\langle f(o)\rangle, P(X)) \rightarrow \neg f(o) |=_p C(S) \tag{5}$$

where symbol $|=_p$ denotes controllability is satisfied when the probability of access actions under control is greater than or equal to *p*.

4. Evaluation Model and Quantitative Analysis of Controllability

To quantitatively analyze controllability of network access actions, we model the network information access control by finite state machine as evaluation model of controllability shown in Fig. 9. The model contains three basic operations of network access control model: information discovery, information analysis and information control.

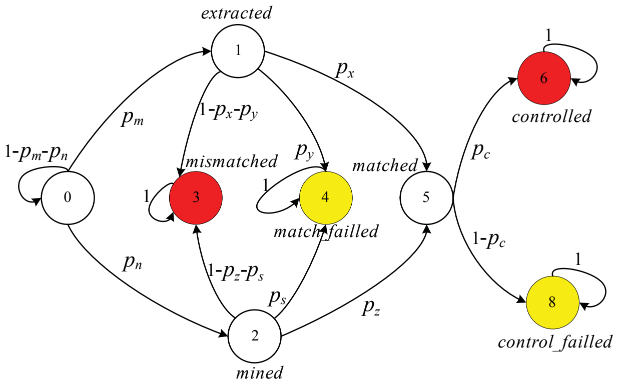

Figure 9. State transition model of network content security access control

Where, state 0 is original state, state 3 and state 6 are controllable end state that denotes information or actions not matching control rules or being successful controlled. State 4 and state 8 are uncontrollable end state that information or actions mis-matching with control rules or being unsuccessful controlled. Arcs denote state transitions by access control processes and symbols p_i on arcs are transition probability. In the model, the transitions *failure* from state 1 or state 2 to state 4 have two possibilities, one possibility is NRM determines actions or information satisfy rules *P*(*X*) but they do not in fact, the other possibility is NRM determines actions or information do not satisfy rules *P*(*X*) but they actually do.

The meanings of the transition probability within evaluation model are shown in Table 1.

Symbol	Meanings
p_m	Probability of information extracted
p_n	Probability of information mined
p_x	Probability of extracted information matched
p_y	Probability of extracted information matched failure
p_z	Probability of mined information matched
p_s	Probability of mined information matched failure
p_c	Probability of information controlled

Table 1. Symbol Description

We adopt symbol Automata to denote the finite state machine shown in Fig.8. There is more than one path to reach uncontrollable end state in Automata. (the controllable end states are state 3 and state 6), so access control model described by Automata is not absolutely controlled. The expression is described below.

$$Automata \;|\neq\; \Box(((extracted \vee mined) \wedge mismatched) \vee \\ ((extracted \vee mined) \wedge matched \rightarrow \Diamond controlled))) \tag{6}$$

Because implementation of information discovery or information analysis or information control is unable to be absolute correct, Internet information access actions can only be controlled in certain probability. Suppose the network access control model fulfills controllability if probability of network access actions controlled is greater than or equal to p, then *Automata* fulfills controllability when $p_m p_x p_c + p_n p_z p_c + p_m(1\text{-}p_x\text{-}p_y) + p_n(1\text{-}p_z\text{-}p_s) \geq p$. Namely,

$$Automata \;|=\; [\Box(((extracted \vee mined) \wedge mismatched) \vee \\ ((extracted \vee mined) \wedge matched \rightarrow \Diamond controlled)))]_{\geq\, p} \tag{7}$$

Control Rate: According to analysis above, we define probability $p_{control} = p_m p_x p_c + p_n p_z p_c + p_m(1\text{-}p_x\text{-}p_y) + p_n(1\text{-}p_z\text{-}p_s)$ as Control Rate of evaluation model. It indicates the probability of information mismatches control rules (the information is permitted) and information matches control rules but is controlled eventually after information is extracted or mined.

Information mining and information extracting are similar. For simple and convenient, we consider the transition probability, which transits to same subsequence state from state 1 and state 2, is equivalent, viz. $p_x = p_z$ and $p_y = p_s$. We also suppose $p_m = 0.8$, $p_n = 0.2$, $p_y = 0$. So the three dimensional graph of Control Rate is illustrated as Fig. 10.

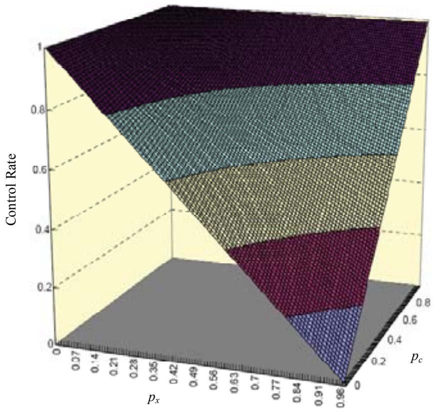

Figure 10. Three Dimensional Graph of Control Rate with p_x and p_c

It is obvious that Control Rate reaches the maximum whatever p_c's value if $p_x = 0$ or whatever p_x's value if $p_c = 1$. That means that information never matches any control rules or information is controlled absolutely if it matches some control rules. Control Rate is 0 when $p_x = 1$ and $p_c = 0$, that means that all information matches some control rules but no one can be controlled.

We suppose $p_m = 0.8$, $p_n = 0.2$, $p_c = 0.9$ in the other case. So the three dimensional graph of Control Rate is illustrated as Fig. 11.

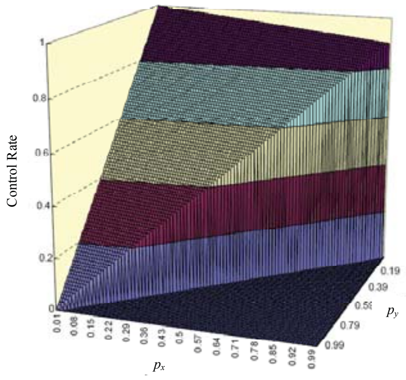

Figure 11. Three Dimensional Graph of Control Rate with p_x and p_y

It is obvious that Control Rate reaches maximum 1 when p_x = 0 and p_y = 0, that means all information is permitted, and reaches minimum 0 when p_y = 1, that means all information has matched wrong. Control Rate decreases from 1 to 0.9 while p_y's value is minimum 0 and increase from 0 to 0.9 as p_x increases.

From analysis above, we consider that match probability p_x and control probability p_c must be done the best to improve in order to increase Control Rate of system in NRM.

Control Leak Rate: The state 6 in *Automata* of Fig.9 is an uncontrollable end state with control leak. We define probability $p_{controlleak} = (p_m \cdot p_x + p_n \cdot p_z) \cdot (1\text{-}p_c)$ as Control Leak Rate of evaluation model. The expression is described below.

$$[\Box(((extracted \vee mined) \rightarrow \Diamond matched) \wedge (matched \rightarrow \Diamond controlled))]_{\geq p} \tag{8}$$

It indicates the probability of information matches control rules but is not controlled eventually after information is extracted or mined. Suppose p_m = 0.8, p_n = 0.2, $p_x = p_z$, the three dimensional graph of Control Leak Rate is illustrated as Fig. 12.

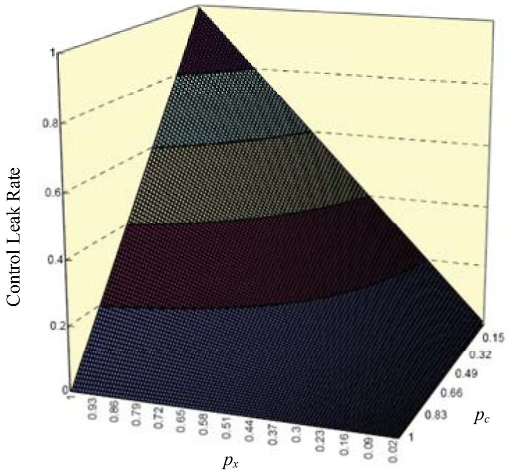

Figure 12. Three Dimensional Graph of Control Leak Rate

Control Leak Rate has maximum 1 when p_x = 1 and p_c = 0, that means all information matched is uncontrollable, and has minimum 0 when p_x = 0 or p_c = 1, that means all information mismatches control rules or matches control rules but is controlled. Control Leak Rate decreases as $p_x \bullet p_c$ increases.

Example 3: Consider a pornographic information control system, which runs in Gateway NRM, all access information and actions are transmitted by the control system. So it made easily to control access actions or information. Assume that the control rules $p(X)$ in the system adopt keyword vectors, URL blacklist and URL white list, where the intersection of URL blacklist and URL white list is empty. The control strategy is as follows, if some action matches some keywords in $p(X)$ and mismatches all URLs in white list in $p(X)$, or it matches some URLs in blacklist in $p(X)$, then the action would be denied. The control system accords with the Internet access control model we present.

In information discovery phase, information of all unencrypted web pages will be extracted for all access information is transmitted by NRM. Probability of $f_{extract}(E,c)$ p_m equals probability of unencrypted web pages. Assume that crack probability of encrypted web information is p_u, probability of $f_{mine}(E,\bar{c})$ is $p_n = (1\text{-}p_m)\, p_u$.

In information analysis phase, the system analyzes information content or web URL with string matching technique. It checks whether the accessed webpage URL matches URL blacklist firstly and suppose match probability is p_v, that is, probability of the webpage belonging to URL blacklist is p_v. Then it analyzes whether the webpage content matches keyword vectors and match probability is p_o, that is, probability of the webpage containing pornographic information is p_o. If the webpage content matches some keyword vectors, it is essential to analyze whether the webpage's URL matches URL white list and suppose match probability is p_w, that is, probability of the webpage containing pornographic information but belonging to URL white list is p_w. Therefore, Probability of $f_{mine}(e,P(X))$ is $p_x = p_v + p_o(1 - p_w)$.

Some errors are occurred while analyzing information. There are two possibilities, one is the webpage is judged matching some rules in $P(X)$ but it mismatches any, the other is the webpage is judged mismatching any rules in $P(X)$ but it matches some. f_{fail}(e, $P(X)$) has two probability, actual error probability p_a and technical error probability p_y. Suppose the only control keyword in $P(X)$ is "oral sex" and the filtered objects are 1000 web pages, where, there are 100 pornographic pages and 900 non-pornographic pages, 300 pages contain keyword "oral sex" while 700 pages do not contain the keyword. As a result of controlling, 305 pages are denied, 295 pages contain the keyword "oral sex" and the other 10 pages do not contain the keyword. Among these 295 pages denied, 95 pages are pornographic, the other 200 pieces are non pornographic. Among the 10 pages denied, 1 page is pornographic, 9 pages are non pornographic. Thus total of 95+1=96 pages containing pornographic information are denied and 200+9=209 pages no containing pornographic information are denied. So the actual match error probability is 1-(95+1)/100+209/900=0.272 and the technical match error probability is 1-295/30+10/900=0.028. Which probability is used depends on the actual application.

The system adopts the same analysis method for mined information, so there is $p_z = p_x$ and $p_s = p_y$.

From the above, probability of $f_{mismatch}$(e, $P(X)$) can be calculated as 1-p_x-p_a or 1-p_x-p_y.

In information analysis phase, probability of $f_{control}(x)$ p_c is up to 0.95 above because the system runs in Gateway NRM and it can strongly control the information user accessed.

Assume that probability of unencrypted webpages is 0.98, crack probability of encrypted webpages $p_u = 0.3$. For all unencrypted webpages can be extracted, so there is $p_m = 0.98$ and $p_n = (1-p_m)\, p_u = 0.006$. Assume match probability of URL blacklist is $p_x = 0.08$ and that of keyword is $p_o = 0.2$, probability of webpages containing pornographic information but belonging to URL white list is $p_w = 0.01$, then probability of f_{match}(e, $P(X)$) is $p_x = p_v + p_o$*(1-p_w) = 0.08+0.2*(1-0.01)=0.278. Assume the technical error probability of f_{fail}(e, $P(X)$) is $p_y = 0.028$, then probability of $f_{mismatch}$(e, $P(X)$) is 1-p_x-p_y = 1-0.278-0.028 = 0.694. And assume probability of $f_{control}(x)$ is $p_c = 0.985$. According to the evaluation model of controllability, probability of Internet information access control is

$$\begin{aligned} p_{control} &= p_m p_x p_c + p_m(1-p_x-p_y) + p_n p_z p_c + p_n(1-p_z-p_s) \\ &= (0.98+0.006)\times(0.278\times0.985+0.694) = 0.954 \end{aligned} \tag{9}$$

And Control Leak Rate is

$$\begin{aligned} p_{controlleak} &= p_m p_x + p_n p_z (1-p_c) \\ &= 0.98\times0.278 + 0.006\times0.278\times(1-0.985) = 0.004 \end{aligned} \tag{10}$$

Assume to use the actual error probability of $f_{fail}(e, P(X))$ p_a and its value is 0.182, and $p_s = p_a$, then probability of Internet information access control is

$$\begin{aligned} p_{control} &= p_m p_x p_c + p_m(1-p_x-p_a) + p_n p_z p_c + p_n(1-p_z-p_s) \\ &= (0.98+0.006)\times(0.278\times0.985+0.54) = 0.803 \end{aligned} \tag{11}$$

If control probability of Internet access or information is required to exceed 0.8, then access of pornographic pages in above example is under control. If control probability of Internet access or information is required to exceed 0.9, then the access of pornographic pages is under control technically, but it is not under control about pages' content in deed. So, we can quantitatively evaluate this control system by way we present. In addition, control efficiency of Internet information access can be enhanced by decreasing match error rate or increasing success rate with different phases in Internet information access control model according to the analysis above.

5. Conclusions and further work

Although information content security is widespread concerned, little work on controllability of content security is studied. We present reference monitor for access control in Internet environment and analyze their advantages and disadvantages firstly. Then we propose definition of network reference monitor and Internet information access control. Abstract operation model for controllability of content security access control are put forward, in the model, Internet access control is divided into three phases: information discovery, information analysis and information control. Then we express intuitive formal specification of controllability by PCTL logic. And we character the process of Internet access control by finite state automata and present evaluation model of controllability to analyze access control of information content security quantitatively.

Controllability of information content security has very important effect in practice, but research on controllability theory is lagging far behind that on controllability application. Although we have made some attempts on controllability theoretical research, it needs further research on controllability model and analysis method to guide practical applications effectively.

Acknowledgment

The work described in this paper was supported by the National Basic Research Program of China (973) (2007CB311100), the National High Technology Research and Development Program of China (863) (2007AA01Z446, 2009AA01Z438) and China Postdoctoral Science Foundation Funded Project (20080440050).

References

[1] Yang, Y. (2005). Summarization of Network Cultural Security. Network Culture and Youth Development Summit Forum.

[2] Stanton, J. J. (2002). Terror in Cyberspace. American Behavioral Scientist 45 (6) 1017-1032.

[3] Goth, G. (2008). Terror on the Internet: a complex issue, and getting harder. *In:* IEEE Distributed Systems Online 9 (3) 3-3.

[4] Volokh, E. (1997). Freedom of speech, shielding children, and transcending balancing. Supreme Court Review 31 (2) 141-197.

[5] Zhang, X. (2006). Communication Decency Act: America's Control Model of Erotic Website. *Journal of Social Sciences,* 8, 136-143.

[6] Qu, Y. W. (2004). Software Behavior. Publishing House of Electronics Industry.

[7] Lin, C., Feng, F., Li, J. (2007). Access Control in New Network Environment. *Journal of Software.* 18 (4) 955-966.

[8] Sandhu, R., Coyne, E. J., Feinstein, H. L. et al. (1996). Role-based access control models. IEEE Computer 29 (2) 38-47.

[9] Park, J., Sandhu, R. (2004). The UCONABC usage control model. ACM Transaction on Information and System Security 7 (1) 128-174.

[10] López, G., Cánovas, O., Gómez, A. et al. (2007). A network access control approach based on the AAA architecture and authorization attributes. *Journal of Network and Computer Applications.* 30 (3) 900-919.

[11] Luo, J., Wang, X., Song, A. (2005). A semantic access control model for grid services. *In:* 9th International Conference on Computer Supported Cooperative Work in Design. IEEE Press, p. 350-355.

[12] Ray, I., Yu, L. (2005). Towards a location-aware role-based access control model. *In:* 1st IEEE International Conference on Security and Privacy for Emerging Areas in Communications Networks. IEEE Press, p. 234-236.

[13] Zhang, H., He, Y., Guo, Z. (2007). An Access Control Model Support Space Xontext. Science in China Series E-Technological Sciences, 37 (2) 254-271.

[14] Park, J., Sandhu, R. (2002). Towards usage control models: beyond traditional access control. *In:* 7th ACM Symposium on Access Control Models and Technologies, p. 57-64.

[15] Jeremy, J. (1988). Security Specification. *In:* Proceedings of IEEE Symposium on Security and Privacy, IEEE Computer Society, Oakland CA, p. 14-23.

[16] Alessandro, A. (2006). Classification of security properties in a Linda-like process algebra. Science of Computer Programming, IEEE Computer Society 63 (1) 16-38.

[17] Jeremy, J. (2005). Logical Omniscience in the Semantics of BAN Logic. *In:* Proceedings of Workshop on Foundations of Computer Security, Stockholm Sweden, p. 1-12.

[18] Zhang, Y., Wang, L., Xiao, G., Wu, J. (2000). Model checking analysis of Needham-Schroeder public-key protocols. *Journal of Software,* Beijing, 11 (10) 1348-1352.

[19] Xu, W., Lu, X. (2003). Model Checking of Authentication Protocols. *Chinese Journal of Computers,* Beijing 26 (2) 195-201.

[20] Ciesinski, F., Groesser, M. (2004). On Probabilistic Computation Tree Logic. Validation of Stochastic Systems, 2925, 147-188.

Authors Biographies

Yin Lihua born in 1973, Ph.D. Her current research interests include computer network and information security, security analysis and intrusion tolerance etc.

Fang Binxing born in 1960, professor, Ph.D. supervisor, member of Chinese Academy of Engineering. His current research interests include computer architecture, computer network and information security etc.

Guo Yunchuan born in 1977, Ph.D. candidate. His research interests include computer network and information security etc.