Security Architectures for Controlled Digital Information Dissemination (CDID)

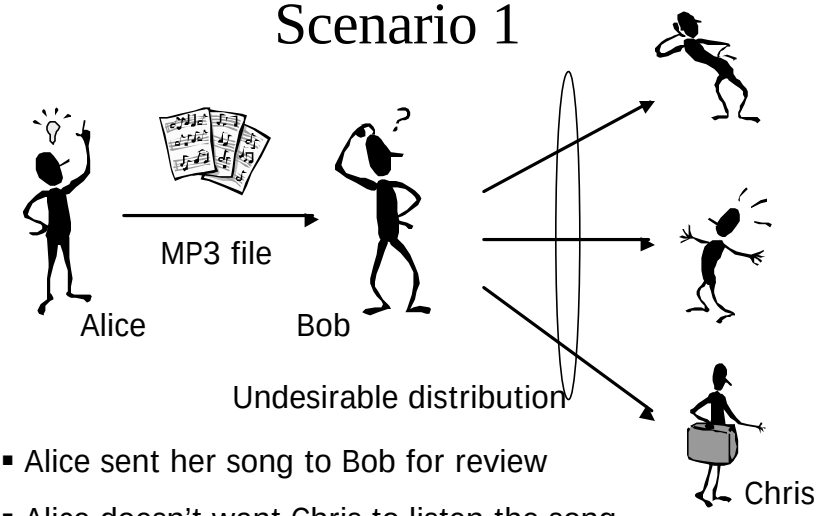
Nov. 2, 2000

Jaehong Park (jaehpark@ise.gmu.edu)
Ravi Sandhu (sandhu@ise.gmu.edu)
James Schifalacqua (JSchifalacqua@si-intl.com)

Overview

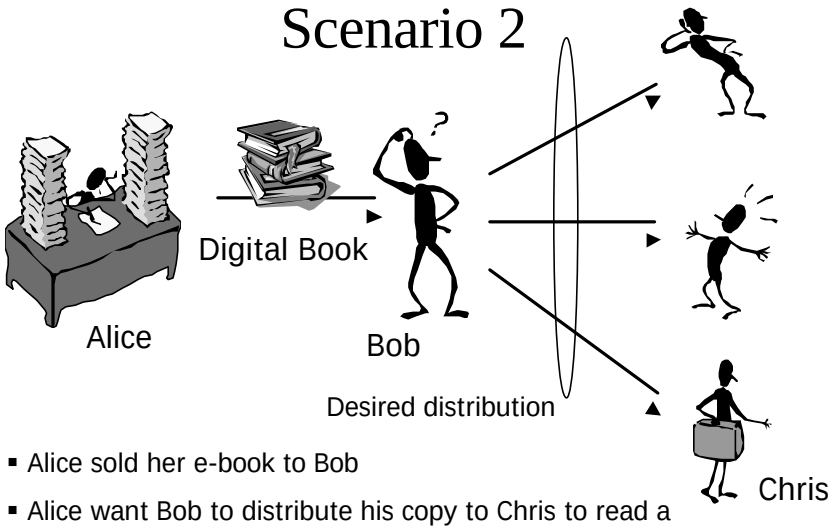
- ***1. Introduction***
- 2. Security Architectures
- 3. Related Mechanisms
- 4. Commercial Examples
- 5. Findings and Conclusions

Scenario 1



- Alice sent her song to Bob for review
- Alice doesn't want Chris to listen the song

Scenario 2



- Alice sold her e-book to Bob
- Alice want Bob to distribute his copy to Chris to read a part of the book so he can be a potential customer

Background

- Unauthorized distribution
  - Reproduction of digital objects does not reduce its quality or value
  - Unauthorized person can access exactly same digital objects as the original copy
  - Commercially, unauthorized dissemination of digital object may cause revenue loss
  - In Intelligence community, unauthorized dissemination causes information leakage

Background (continued)

- Superdistribution
  - First introduced by Dr. Mori in 1983
  - Digital information is wrapped with digital strings and freely available to everyone
  - No restrictions on copying
  - Copying and distribution is encouraged for marketing purposes
  - Digital information is accessible only where special software is available
  - The usage is controlled by appropriate authorities
  - Well accepted in portions of commercial world

Background (continued)

- Application level security solutions
  - Lower layer security cannot support controlled dissemination
  - Application level security solutions can be used
  - Securing digital object itself, not the transmission
    - By using cryptographic, watermarking, or use-control technologies

Objectives

- To develop application-level security solutions that secure digital information itself for controlled dissemination (rather than securing transmission of digital information at lower layers)
- To develop systematic security architectures for controlling and tracking digital information dissemination

Dissemination Attributes

- Dissemination Scale
  - Small, medium, and large scale
- Dissemination Environment
  - Closed, federated, and open environment
- Payment-based vs. Payment-free
- Prevention vs. Detection & Tracking

Dissemination Scale

- Small Scale Dissemination (SSD)
  - 1 item → 1 to 100 recipients
  - Much less tolerance for leakage
  - B2B Business transaction, Intelligence community
- Medium Scale Dissemination (MSD)
  - 1 item → 10^3 to 10^5 recipients
  - Textbook publishing, technical journals
- Large Scale Dissemination (LSD)
  - 1 item → 10^6 to 10^8 recipients
  - Some leakage is acceptable or even desirable
  - Music, popular books

Dissemination Environment

- Closed Environment Dissemination (CED)
  - Internal distribution (commercial and Intelligence)
  - Easy to customize Client-side systems (both S/W & H/W)
- Federated Environment Dissemination (FED)
  - Limited number of organizations are involved
  - B2B, B2G and G2G dissemination
  - Limited administrative control over recipients
- Open Environment Dissemination (OED)
  - B2B and B2C dissemination
  - Hard to customize client-side system

Two Types of Dissemination

- Payment-Based Type (PBT)
  - Payment is required in order to access digital content
  - B2C mass distribution e-commerce system
- Payment-Free Type (PFT)
  - Payment is not required
  - Dissemination must be controlled for confidentiality or other security requirements
  - B2B Hub System, Intelligence Community

Characteristics of PBT & PFT

- PBT
  - A small amount of information leakage is acceptable and even desired
  - The number of legitimate copies of a single digital item is usually greater than that of PFT
  - The objective in PBT is to distribute as many copies as possible and extract payment
- PFT
  - Information leakage is not acceptable
  - The number of legitimate copies of a single digital item is less than that of PBT
  - It is the distribution itself which needs to be limited
  - The security requirements are likely to be more stringent than that of PBT

Prevention vs. Detection

- Prevention
  - Leakage “cannot” occur
- Detection & Tracking
  - Leakage can occur but can be detected and tracked to re-distributor
- Both solutions can/must coexist

Commercial Interest

	Payment	Scale	Environment	Protect vs. Detect
Major Commercial Interest	Yes	Large Medium	Open Federated Closed	Both
Less Commercial Interest	No	Medium Small	Open Federated Closed	Both (Prevention emphasis)

Scope

- We are focusing on PFT.
  - Control dissemination solutions of PBT have been developed actively in commercial sector
  - However, no systematic study for more generalized security architectures for controlled digital information dissemination has been done
  - Architectures can be extended to include payment function

Overview

- 1. Introduction
- ***2. Security Architectures***
- 3. Related Mechanisms
- 4. Commercial Examples
- 5. Findings and Conclusions

Three Factors of Security Architectures

- Security Architectures have been developed based on the following three factors
- Three factors
  - Virtual Machine (VM)
  - Control Set (CS)
  - Distribution Style

Three Factors of Security Architectures (continued)

- Virtual Machine (VM)
  - A module that runs on top of vulnerable computing environment and has control functions to provide the means to control and manage access and usage of digital information
  - Foundation of use-control technologies
  - Needs for specialized (trusted) client software/hardware

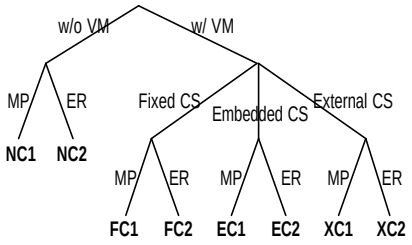
Three Factors of Security Architectures (continued)

- Control Set (CS)
  - A list of access rights and usage rules that is used by the virtual machine to control a recipient's access and usage of digital information
    - A *fixed control set* is hardwired into the virtual machine
    - An *embedded control set* is bound to each digital object
    - An *external control set* is separate and independent from the digital object

Three Factors of Security Architectures (continued)

- Distribution Style
  - Message Push (MP) style
    - Digital information is sent to each recipient
  - External Repository (ER) style
    - Each recipient obtains the digital information from dissemination server on the network

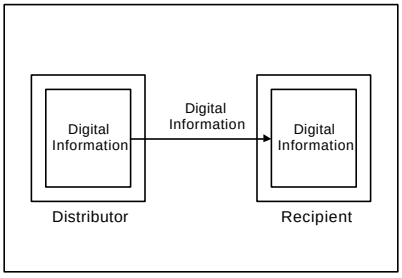
Architecture Taxonomy



VM: Virtual Machine
CS: Control Set
MP: Message Push
ER: External Repository

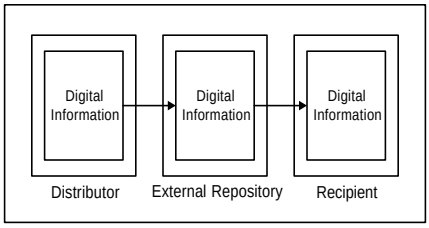
NC1: No control architecture w/ MP
NC2: No control architecture w/ ER
FC1: Fixed control architecture w/ MP
FC2: Fixed control architecture w/ ER
EC1: Embedded control architecture w/ MP
EC2: Embedded control architecture w/ ER
XC1: External control architecture w/ MP
XC2: External control architecture w/ ER

No Control Architecture w/ Message Push (NC1)



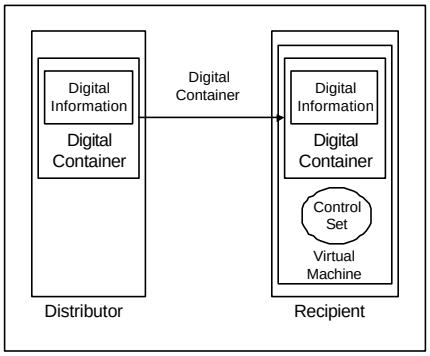
- Distributor directly sends a copy of digital contents to each recipient
- Each recipients stores the copy of digital information at local storage
- After distribution, no direct means to control the distributed digital information
- To access the digital information from multiple system, the recipient needs to transport the information

No Control Architecture w/ External Repository (NC2)



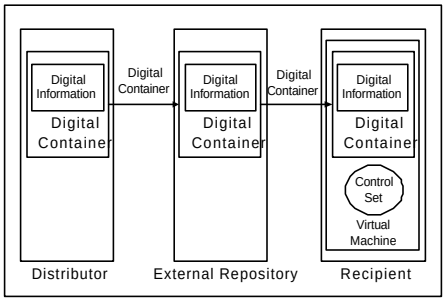
- Digital information is sent to an external repository server for distribution
- A recipient must connect to the external repository to access the digital content
- Once a recipient has received the digital contents, there is no way to control access or usage

Fixed Control Architecture w/ Message Push (FC1)



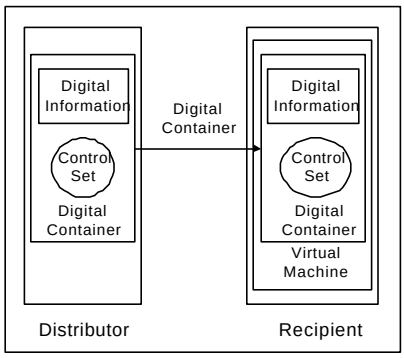
- Digital content is encapsulated in a digital container
- Control set is encoded into virtual machine
- The control set cannot be changed after the distribution of the virtual machine
- Access is controlled based on control set
- Each recipient should keep the received information for further access to it

Fixed Control Architecture w/ External Repository (FC2)



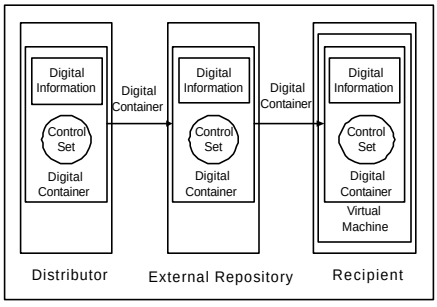
- Similar to FC1, except that digital container is sent to external repository for distribution
- A recipient must connect to the external repository to access or download the digital container
- Accessibility to the content by a single recipient from multiple computers

Embedded Control Architecture w/ Message Push (EC1)



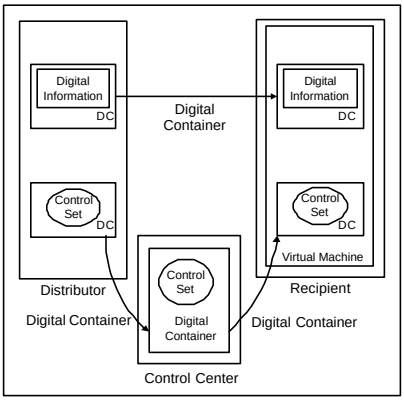
- Control set is embedded in the digital container with digital information
- Distributed content will be controlled based only on the pre-set access rights and usage rules
- After distribution, distributor cannot change the control set of the distributed digital content
- Recipients can access digital content without any network connection
- Only pre-set revocation is available

Embedded Control Architecture w/ External Repository (EC2)



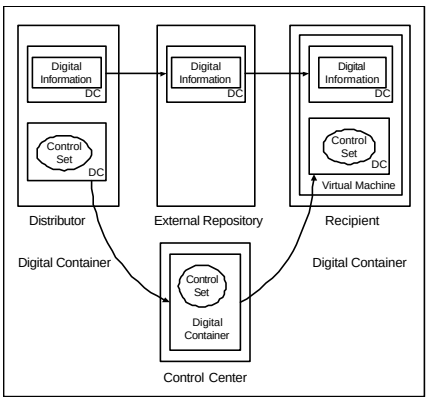
- Digital container is sent to the external repository server for distribution
- If digital container is prohibited from being locally stored, the distributor can revoke a previous granted access by changing control set

External Control Architecture w/ Message Push (XC1)



- Control set can be encapsulated independently from digital content
- Two possible options:
  - Network connection is always required
  - Network connection is required from time to time (one time connection is possible)

External Control Architecture w/ External Repository (XC2)



- Separation of content and access rights
- 4 variations
  - Both encapsulated digital content and encapsulated control set can be stored on recipient's local storage
  - Encapsulated digital content is freely available, but control set cannot be locally stored
  - Only encapsulated control set can be stored
  - Neither can be stored locally

Security Characteristics

	Characteristics	NC1	NC2	FC1	FC2	EC1	EC2	XC1	XC2
C1	Disseminator can control access and usage of disseminated digital information			Y	Y	Y	Y	Y	Y
C2	Disseminator can change recipients' access rights after dissemination						Y	Y	Y
C3	Re-disseminated digital information can be protected			Y	Y	Y	Y	Y	Y
C4	Special client software (virtual machine) is vulnerable to attacks			Y	Y	Y	Y	Y	Y
C5	Tracking re-disseminated digital information is possible	Y	Y	Y	Y	Y	Y	Y	Y

Functional Characteristics

	Characteristics	NC1	NC2	FC1	FC2	EC1	EC2	XC1	XC2
C6	Disseminated digital container is reusable for other recipients by re-dissemination							Y	Y
C7	Digital information does not have to be on recipient's storage		Y		Y		Y		Y
C8	Digital information can be accessible from any machine if it is connected to network		Y		Y		Y		Y
C9	Recipient should carry digital information to access it from multiple machines	Y		Y		Y		Y	
C10	Special client software (virtual machine) is required			Y	Y	Y	Y	Y	Y
C11	In case of large digital information, download time can be significantly costly		Y		Y		Y		Y
C12	Every access to digital information requires network connection.								
C13	The architecture can be supported without network connection	Y		Y		Y			
C14	Control center trusted by both distributors and recipients is mandatory							Y	Y

Commercial Solutions

Solution	Organization	NC1	NC2	FC1	FC2	EC1	EC2	XC1	XC2
Adobe Acrobat	Adobe					X			
PDF Merchant & WebBuy	Adobe								X
PageVault	Authentica							X	
SoftSEAL	Breaker Technologies								X
Confidential Courier	Digital Delivery, Inc.					X			
docSPACE	DocSPACE Co.		X						
CIPRESS	Fraunhofer Institute for Computer Graphics & Mitsubishi Co.								X
Cryptolope	IBM							X	
InTether	Infraworks Co.					X			
InterTrust	InterTrust Technologies Co.							X	
RightMarket	RightMarket.com Inc.							X	

Overview

- 1. Introduction
- 2. Security Architectures
- ***3. Related Mechanisms***
- 4. Commercial Examples
- 5. Findings and Conclusions

Digital Watermarking

- Digital Watermark
  - Digital watermark is used to mark the identity of the objects with information such as author's name, date, or usage right
  - Can provide tracking capability to illicit distribution
  - Can be implemented all of our security architectures
  - Watermarking technologies are dependent on the type of digital information (e.g., text, image, audio and etc.)
  - Minimum size of object is required
  - Difficulty of embedding different watermark (fingerprint) in each copy of original objects in case of mass distribution

Digital Watermarking

- Visible Watermarking
  - Watermark is visible (e.g.background logo)
  - Can be used for sample digital objects to reduce commercial value on them
- Invisible Watermarking
  - Most of Watermarking belongs here
  - Digital Watermark can be detected by special software

Digital watermarking

- Public Watermarking
  - Watermark information w/ publicly known key (w/o any secret key)
  - Everyone can read watermarked information
  - In commercial sector, customer can find copyright owner's information
- Private Watermarking
  - Only authorized users can detect the watermarks
  - Good for tracking purpose

Digital Watermarking

- Watermark retrieval
  - Reference-required watermark
    - The original object or embedded watermark information is required for comparison
  - Reference-free watermark
    - Watermark can be retrieved without the original document or added information
    - The mechanism detects specific properties and patterns from documents

Digital Watermarking

- Different format, different watermarking
  - text, image, audio, and video
- Text Watermarking (Brassil, et al.)
  - Line-shift coding
    - Text lines are shifted imperceptibly up and down
  - Word-shift coding
    - Words are shifted horizontally
    - Original un-watermarked documents are required for extracting watermarked information
  - Feature coding
    - Characters are altered (vertical or horizontal)
    - Least discernible, larger information embedded

Use-Control Technologies

- Use-Control technologies
  - Originally based on Superdistribution concept
  - Digital information is encapsulated into a cryptographically protected electronic container
  - Digital information is only accessible by using a special application software with appropriate access rights
  - A Control Center exists for controlling and managing the access rights, usage rules, and even usage history

Overview

- 1. Introduction
- 2. Security Architectures
- 3. Related Mechanisms
- ***4. Commercial Examples***
- 5. Findings and Conclusions

Adobe PDF Merchant & WebBuy plug-ins

- PDF Merchant
  - Server-side software that encrypts PDF file, generates title-key and distributes the license files (Vouchers) with the title key
  - Security options (disable print, copy, change, or annotate function)
  - Binding contents to either CPU-ID, network ID, e-mail address, time or portable media

Adobe PDF Merchant & WebBuy plug-ins (continued)

- WebBuy
  - A plug-in software within Acrobat Reader 4.05
  - Our Virtual Machine (VM)
  - Controls access to PDF files by using Voucher information

Adobe PDF Merchant & WebBuy plug-ins (continued)

- Vouchers
  - Is generated using Merchant when a customer buys rights to use PDF files
  - Grants access rights to the customer
  - Signed XML 1.0 files
  - Include title key to decrypt locked PDF file
  - Include access rights and usage rules
  - 1,024 RSA digital signature

Commercial Efforts for Open-standard (XrML)

XrML: Extensible Rights Markup Language

- What is XrML?
  - "A language in XML for describing specifications of rights, fees and conditions for using digital contents, together with message integrity and entity authentication within these specifications"
  - An extension of the Xerox "Digital Property Rights Language version 2.0 (DPRL)"
  - ContentGuard™ has developed XrML as an open specification licenced on a royalty-free basis
- Why XrML?
  - In CDID Architecture, XrML can be viewed as one of potential mechanisms for Control Set (CS) implementation.
  - XrML is extensible, open specification

XrML in Digital Right Management (DRM)

Source: www.xrml.org

Top-Level Structure

```
<XrML>
    <BODY>
        (TIME)?                      time interval in which this spec is valid
        (ISSUED)?                    time moment at which this spec is issued
        (DESCRIPTOR)?                description or meta data of this spec
        (ISSUER)?                    principal who issues this spec
        (ISSUEDPRINCIPALS)?          list of principals this spec is issued to
          (PRINCIPAL)+
        (WORK)?                      work and rights this spec specifies
        (AUTHENTICATEDDATA)?         data that provided to application
    </BODY>
    (SIGNATURE)?
</XrML>
```

"?" denotes zero or one occurrence; "+" denotes one or more occurrences; and "*" denotes zero or more occurrences

Digital Works & Usage Rights

```
<WORK>
    OBJECT          object used to identify the work
    DESCRIPTION     description of the work
    CREATOR         creator of the work
    OWNER           owner of the work
    METADATA        additional metadata of the work
    DIGEST          digest value of the work
    PARTS           parts of the work, each of which is a work itself
    CONTENTS        indicator of where content of the work is
    COPIES          number of copies of the work that are specified
    COMMENT         Comment
    SKU             Stock Keeping Unit, for extensibility to allow people to
                    identify this work in their own stock.
    FORMAT          Digital or physical manifestation of the work
    (RIGHTSGROUP | REFERENCEDRIGHTSGROUP )+   rights group associated
                    with the work | reference rights group of the work
</WORK>
```

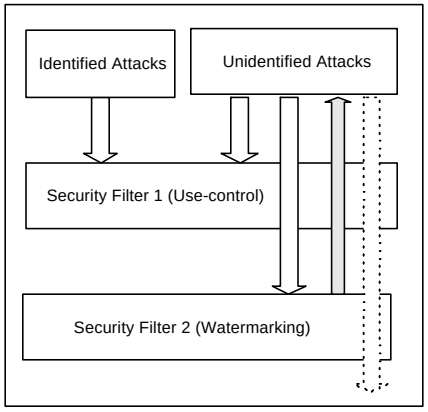
Rights in RIGHTSGROUP Element

- Digital property rights
- Specifying times, fees and incentives
- Specifying access controls (licenses/certificates, security levels)
- Specifying territory information
- Specifying tracking information
- Specifying watermark information
- Bundle specifications (time limits, fees, access, and watermark info inside bundle)

Overview

- 1. Introduction
- 2. Security Architectures
- 3. Related Mechanisms
- 4. Commercial Examples
- ***5. Findings and Conclusions***

Solution Approaches



- Two layer approach (protection and tracking)
- Use-control mechanism can provide a certain level of protection from attack
- Watermarking mechanism gives tracking ability
- Watermarking technologies are still premature to guarantee tracking of dissemination

Conclusion

- Layered approach: security objectives, security architectures, and mechanisms
- The first systematic study on security architectures which are not previously defined in this manner
- Our work provides basis for future researches and development for controlling and tracking dissemination of digital contents

Future Researches

- Studies on mechanisms
  - Use-control (e.g., virtual machine, control set)
  - Watermarking technologies
- Studies on security architectures in detail