Controlling Primary And Secondary Access To Digital Information

Marshall D. Abrams
The MITRE Corporation, 1820 Dolley Madison Blvd.,
McLean, VA 22102, abrams@mitre.org
voice: 703-883-6938 fax: 703-883-1397

Paul B. Schneck
MRJ Technology Solutions, 10560 Arrowhead Drive
Fairfax , VA 22030-7305, schneck@mrj.com
voice: 703-277-1618 fax: 703-277-1701

Abstract

A system is described for controlling primary and secondary access to digital information when the recipient is not fully trusted. Tamper–detecting hardware implements the controls. Cryptography protects both information and access rules in storage and during distribution. The system is applicable to national security and commercial models, including entertainment and electronic commerce.

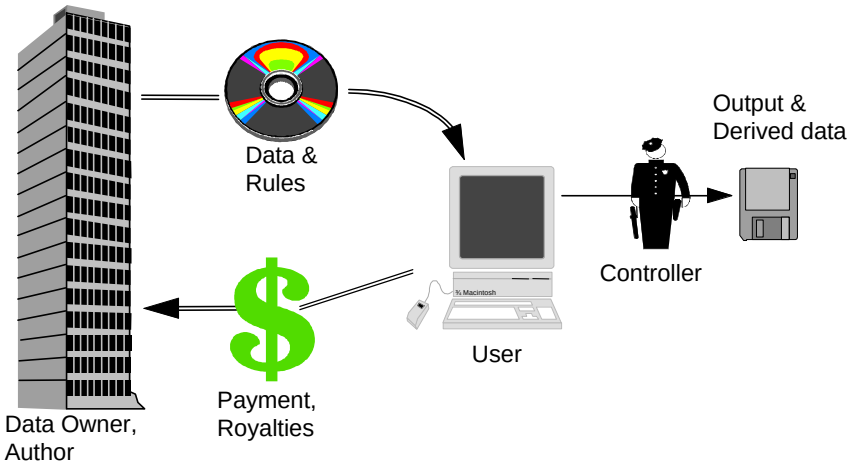

Figure 1. Controlling Access to Content

Overview

This paper describes a system* (US patent 5,933,498; additional US and foreign patents pending) called Persistent Access Control (PAC), for controlling access to data. PAC is intended for use in environments where the user is not trusted, in whole or under some conditions. Some of the high-profile environments that PAC is applicable to include national and industrial security, electronic commerce, and entertainment. A high-level representation of a common application of PAC, Figure 1, shows the author distributing digitized content on CD-ROM. Processing on the computer and output are controlled. Required payments are made to the author. The access control policy is expressed in a set of rules. Users

* This work was performed under sponsorship of The MITRE Corporation, where both authors were then employed.

obtain access to the content only in accordance with the rules, which are enforced by a mechanism protected by tamper detection. For another view of the PAC concepts see [9]. PAC provides for:

- Controlling distribution of data for subsequent use
- Protecting all or selected portions of the data
- Preventing access to the protected portions of the data (other than in a non-useable encrypted form)
- Determining rules concerning access rights to the data
- Protecting the rules
- Packaging the protected portions of the data and the protected rules
- Preventing access to protected rules
- Permitting access to the protected data only in accordance with the rules

Background

Motivation

Commercial and national security information is inadequately protected against individuals abusing their rights of access. One reason that the "insider problem" has not received adequate attention is that no adequate technology has been available to help control access rights. Actually the simple distinction between insiders and outsiders is inadequate; most users should have limited rights. Access control policy models include need-to-know, Chinese wall, labels, and role-based.

The Software Publishers Association and the Business Software Alliance estimate that software worth billions of dollars is illicitly copied each year. Print and entertainment publishers hesitate to expand into the Internet marketplace because they are unable to control (in the sense of receiving compensation in return for rights) secondary distribution of their products or incorporation of their products into derivative products. As discussed in "Rules and Policies" below, the rules can distinguish many classes of users if the policy requires such distinction.

PAC is a computerized implementation and extension of concepts that originated in originator controlled (ORCON) [4] data dissemination–a manual control method for paper documents. ORCON policy requires the permission of the originator to distribute protected information to personnel not originally designated as authorized recipients by the originator. Previous research addressing the automation of ORCON policy includes [1, 2, 5, 6, 7, 8].

Existing Technology

The principal technology used for controlling the distribution of digital information is cryptography. The information to be protected is encrypted and transmitted to the authorized user(s). Separately, a decryption key is provided only to authorized users. The key is subsequently used to enable decryption of the information so that it is available to the authorized users. Cryptography cannot protect the data after decryption. Thus, secondary distribution and multiple inappropriate uses are possible. Access control mechanisms, used when the information is made available to the authorized user, cannot control the information in the hands of others.

Older controls–including tokens, dongles, so-called “uncopyable” media, various executable software protection schemes, and executable software for printing that places an identifier on all printed output in a fashion not apparent to a human–fail to limit secondary distribution or distribution of derivative works.

This shortcoming is not a failure of mechanism, but rather is an architectural design omission. The problem of copying by the authorized user is simply not addressed. Once data are available to an authorized user, they are uncontrolled and may be copied, modified, or transmitted at will. Identifiers can be included on printed output to help identify the source of copied material, but they do not prevent secondary distribution.

Threats

Some common threats to data and processing systems include the following:

- Willful or accidental violation of policy
- Digital copying
- Capture of output signal
- Deliberate attack via legacy or customized hardware

Policy Violation. In national security environments policy violation is the major consideration. In commercial models lack of compensation to the owner is a significant disincentive to provide the program material in digital form.

Digital Copying. Once data are decrypted, the resulting cleartext must be protected from unauthorized copying. Creating an unauthorized local copy or disseminating the data without authorization results in an original-quality copy which can be distributed in violation of security policy. Indeed, as noted above, a digital copy is identical to the original. The concept of “copy” loses its significance in this context.

Capture of Output Signal. No matter what method is used to protect a file, its data can be captured as a signal en route to an output device, such as a display or printer. Capture of an analog signal results in some degradation of signal quality. However, the market for bootleg copies of videos, for example, appears to be insensitive to reduced quality if the price is right. A captured digital signal suffers degradation of quality only as a result of bit errors (i.e., if the data capture was not completely accurate). Bit error rates of 10^{-9} are common. For all practical purposes, digital copies are identical to the original.

This threat is well known to the entertainment industry. Ciciora [3] discusses various approaches to protection that have been incorporated in TV set-top boxes. An early example is that set-top boxes, when tampering was detected, would put “jitter” and/or “snow” on the picture. When the subscriber called for service, the company alerted by the symptoms, would check the set top box, and know that the subscriber had tampered!

Deliberate Attack via Legacy or Customized Hardware. High-intensity attack by attackers possessing a high level of expertise, opportunity, resources, and motivation must be considered. Such attackers might include foreign governments, industrial espionage agents, and resellers of pirated digital information. This threat exists in uncontrolled hardware. For example, the inadequately protected information would be available in the memory and could be accessed via dual-ported memory or even by direct memory access from a peripheral device.

PAC Capabilities

PAC controls access, use, and distribution of data. For example, when the content is in the form of textual and graphical information, PAC can control how much of the information is displayed and in what form. When the content is a computer software program, PAC can control how much of the software's functionality is available. PAC allows users to buy or “rent” features.

Systems that rely only on software for security can be defeated. Software mechanisms can be bypassed or rendered ineffective if access to the hardware is not controlled. Modifying the software by removing the code that inhibits access, modification, or output renders access control ineffective. Similarly “snooping” on the bus or unloading memory results in a copy of the cleartext. PAC protects the hardware by detecting tampering and rendering access impossible by destroying the data themselves as well as critical access control data. Without the tamper detection/reset mechanism of PAC, software can be modified or data can be intercepted, thereby rendering useless any attempts at control. Like most algorithmic systems, PAC can be implemented in software or firmware, protected by the tamper detection/reset mechanism to prevent subversion. For increased performance, the algorithms can be implemented in hardware.

PAC relies on tamper detection to provide sufficient advance notice to permit destruction of data and cryptographic variables. Tamper detection is a tractable problem that should not be confused with tamper–proofing. Several techniques for tamper–detecting packaging are described in [10]. In PAC, if and when tampering is detected, at least the following operations are performed. The cryptographic variables (e.g., keys) are destroyed, all rules are destroyed, all cleartext information is destroyed, all files are closed, and the device is otherwise deactivated. While these operations are described sequentially, in practice they occur simultaneously or in some concurrent or parallel order. If some order must be imposed on these operations, the first priority is to erase the cryptographic variables. Tamper detection needs only a microsecond to wipe out the cryptographic variables (by active rewriting); milliseconds to erase the RAM.

The rules may be packaged with the content or may be provided separately. The rules specify the access rights granted to the user, including rights of further distribution of the content. PAC can be implemented in a stand-alone device such as a television, a VCR, a laser printer, a telephone, or a computer system. The rules, policies and protections of content are typically made by the content owners and/or distributors based on their analysis of applicable security threats. National security “mandatory” policies can be installed by equipment owners (e.g., the government).

Threats & Countermeasures

This section provides a few examples of threats and countermeasures.

Threat: Capture of Output Signal

Countermeasure 1: Encrypt or Scramble Output Signal

Protection of the output signal is accomplished with encryption of a digital signal and scrambling of an analog signal. The information must be protected within the output device. This solution requires installing decryption or unscrambling capability in the output device, TV, or monitor, along with an appropriate tamper detection capability. Encryption or scrambling might be effected using a public key associated with the output device (although, to prevent so-called "spoofing," the key should be obtained from or validated by a certification authority and not from the output device). Alternatively, the output might be encrypted or scrambled using a

private key dedicated to the designated output device. The output signal is decrypted or unscrambled by the output device using its private key and is not available in the clear outside of the device's protected enclosure. For example, the “5C” system, a proprietary interface developed by 5 Companies—Intel, et al, negotiates a session key.

Countermeasure 2: Protect Output Signal by Packaging

The output signal is protected by making it unavailable outside the device. Examples of such packaging include lap-top computers and the all-in-one Macintosh computer, as well as integrated televisions, VCRs and video or audio laser disk players. A sealed-unit computer with tamper detection would provide the necessary protection.

Threat: Digital Copying

Countermeasure 1: Secure Coprocessor

Use of a secure coprocessor is indicated to protect against unauthorized copying [10] when an operating system (OS) is untrustworthy—that is, when an OS cannot provide adequate resistance to the anticipated threat. When the OS is untrustworthy, any countermeasures implemented in the OS, or protected by it, can be circumvented through the OS or through bypassing it. In contrast, a trustworthy coprocessor can provide adequate security functions with known assurance.

Countermeasure 2: Detection of Unsealing

Sealing is necessary but not sufficient. Detecting of unsealing and appropriate erasure of critical data is also required. In PAC the system is protected by tamper detection, which causes the rules, cryptographic data, and decrypted protected data to be destroyed when tampering is detected. Both passive and active means are used to effect such destruction. Semiconductor memory is volatile and does not retain data when power is removed. A long-life battery provides energy sufficient to allow rewriting (zeroizing) nonvolatile memory containing, for example, the private key, without which the system will be unable to decrypt any protected data.

Using Cryptography

Authoring refers to selecting the portions of the content to be protected and packaging the content, rules, and keys. Asymmetric encryption algorithms are employed in authoring and access control. The keys for these algorithms are protected within the PAC-protected system and are never exposed. The data-encrypting key, K_D, selected by the distributor, may be the same for all copies of a specific data package. K_D may be different for each distinct data package. The symmetric encryption algorithm used for encrypting the content is associated with K_D and may also be selected by the distributor. K_D is encrypted using a rule-encrypting key K_R. When the rules are distributed with the packaged data, K_R may be the same for all packages. When the rules are distributed separately from the packaged content, K_R can be unique for each version of the PAC-protected system. The rule-encrypting key K_R is known only to (and protected within) each receiving computer of each user.

Figure 2 shows a flow chart of a version of the authoring process in which the rules are distributed with the packaged content, the distributor (acting as a representative of the owner of the data) selects a content-encrypting algorithm and K_D, and then encrypts K_D using K_R. The encrypted K_D is then stored in the encrypted ancillary information of the packaged data.

The algorithm selection is based on an assessment of risk, the degree of protection desired and other relevant factors, such as speed, reliability, and exportability. *Risk* refers to the expected impact of anticipated threats.

K_D may be generated in a manner suitable for the selected data-encrypting algorithm. For data having lower value to its owner, or having lower risk, all distributions may rely on a small number of data-encrypting keys. Another encryption method uses a unique K_D for each item of content to be distributed.

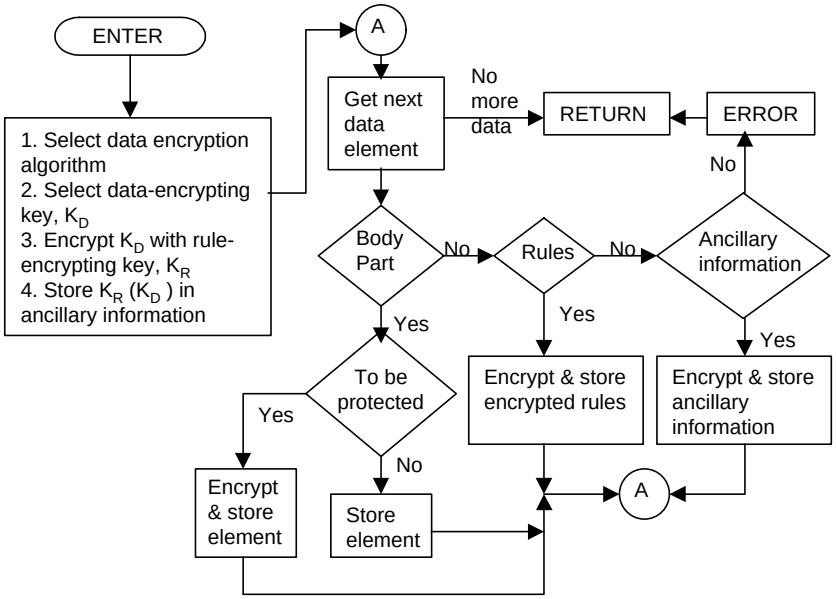

Figure 2. Flow Chart Of A Version Of The Authoring Process

Having selected a content-encrypting algorithm and K_D, and having encrypted and stored K_D, the distributor processes the various elements of the data. The rules are protected by encrypting them using K_R, and the encrypted rules are then stored in the encrypted rules part of the packaged content.

Authoring

The granularity depends on the type of restrictions needed on content use and on the form in which the content is provided. The distributor obtains and examines each content element at the desired granularity and determines whether the element being examined is in the body of the content (as opposed to being rules or ancillary information).

If the element being examined is determined to be in the body of the data, the distributor then decides whether the element is to be protected, that is, whether or not access to that element of the content is to be controlled and the content element is to be encrypted. For example, if the content represents a textual article, the article abstract might be unprotected even though the body of the article is encrypted.

If the current content element is determined to be ancillary information, the ancillary information is protected by encrypting it using K_D and then the encrypted ancillary information is stored in the encrypted ancillary information part of the packaged data.

A hybrid implementation, wherein some rules are packaged with the content and other rules are packaged separately, is also possible.

Self-Protecting Packaging

All components of the protected content are packaged in such a way as to exclude any unauthorized access by a user. That is, the access mechanism is packaged in a tamper–detecting manner. Once tampering is detected, the access mechanism is disabled. All components of the tamper detection mechanism are part of the access mechanism, connected to processing unit, energy source, and non-volatile memory.

PAC employs a combination of physical self-protection measures coupled with means for detecting protection violations. Critical data are destroyed when a violation is detected. For example, hardware circuits can accomplish the following essentially simultaneously:

- Make the access mechanism inoperative
- Destroy (zeroize) all cryptographic keys and data
- Clear non-volatile memory

Tamper detection allows the access mechanism to ensure that all internal data (both the system's data and any user data) are destroyed before any tamperer can obtain them.

One way to deny access to the data within the access mechanism is to package all of the components within a single physical case. For example, a typical portable laptop computer meets the requirement of having all components within the same physical package or case. Detection that a case has been opened is straightforward, and the detection method is well known.

All content stored on non-volatile memory units, e.g., hard disk, is encrypted. The encryption of the content stored on the hard disk can use cryptographic keys generated within the access mechanism, which are never known outside the mechanism. In this way, when tampering is detected, the cryptographic keys are lost. Therefore, if the hard disk is removed from the mechanism, any data stored thereon will be inaccessible without the appropriate keys. This is necessary because, unlike volatile media, a disk cannot be erased quickly when tampering is detected.

In general, within the PAC-protected system, the data are encrypted on any non-volatile storage devices so that they remain unavailable in the case of tampering—specifically, a “sledge hammer attack” designed to destroy any mechanism that might erase protected (but unencrypted) content. Unencrypted content (1) is only present within the access mechanism (2) inside the security boundary (3) in components where the content can be immediately destroyed when tampering with the access mechanism is detected.

Operational Considerations

Certain operational procedures may be required to maintain the protections and controls. Operational procedures may be employed to prevent the production of equipment that includes PAC concepts and contains methods for circumventing protections and controls.

These operational procedures involve inspection, analysis, testing, and perhaps other procedures, followed by certification of authorized access mechanism implementations. The inspection might include design analysis and physical chip inspection. Upon successful

inspection, a cryptographically sealed certificate is stored within the protection perimeter. This certificate is one of the data items destroyed upon detection of tampering. The certificate is issued by an authorized Certificate Authority (CA) and includes a decryption key issued by that CA.

K_R may be encrypted using the encryption key corresponding to the decryption key included in the certificate in each device. In order to obtain K_R within the device, the device must have the decryption key that was stored in the certificate by the CA.

Reverse Path Transactions

Some data may flow from the user back to the data owner. Acknowledgement of receipt may be provided and protected by non-repudiation cryptography. Receipts can be used for document tracking, reminder of suspense dates for reviews, and auditing.

In our market economy, producers and distributors of goods and services expect to be compensated. Digital information producers and distributors are no exception. The needs of commerce have been a primary factor in the evolution of information technology throughout history. Many of today's information infrastructure activities also deal with billing and payment.

PAC can employ an appropriate electronic payment system. There are many competing electronic payment systems available. Some operate in real time by communicating through the Internet or direct dial. Others employ a prepaid balance which is debited against merchant credits, with periodic batch updating and transmission.

Rules and Policies

The rules (provided together with or separately from the packaged content) embody the content-owner's policies for controlling users' access rights to the protected data.

PAC permits the owner of digital information to sell or license various levels of access rights to the protected data and to ensure no access beyond those rights. PAC allows only the type and quantity of access approved by the owner.

While the rules allowed are open-ended, an example set of access control parameters is given below. Access control parameters may be combined to enforce arbitrarily complex policies. Some parameters are independent of any other parameters, some parameters are mutually exclusive, and some parameters must be used in combination to define fully the actions to be allowed or disallowed.

No restriction

This would be the status if no restrictions were placed on the associated content. If this parameter is explicitly stated, it overrides any contradictory parameter that may also be present. The content may be read, printed, executed, modified and copied.

No Modify

The content may not be edited or changed.

No Copy

The content may not be copied, and a derivative work may not be made from the content.

No Execute

The content may not be executed.

No Print

The content may not be printed.

Print With Restriction of Type n

If the user prints after accessing the content, a simulated watermark will be printed as background, or a header and/or footer will be placed on each page. The numeral *n* specifies the specific restriction to be applied, e.g., standard watermark (such as “do not copy”), personal watermark (such as “printed for *name of user*”), standard header/footer (such as “*Company Name* Confidential”), or personal header/footer (such as “Printed for *name of user*”).

No Access

Any user access, including an attempt to execute, will retrieve only encrypted content (ciphertext). This is the default case when there are no rules associated with content or when the rules are corrupted.

No Child Access

Unless the user has been identified as an adult (for example by use of a password or a token) access will not be allowed for items identified as “adult material.”

Access Cost = (unit, price)

Each time a *unit* of content (e.g., book, volume, chapter, page, paragraph, word, map, record, song, image, kilobyte, etc.) is opened, a cost of *price* is incurred.

Print Cost = (unit, price)

Each time a *unit* (e.g., page, file, image) is printed, a cost of *price* is incurred.

Copy/Transmit Cost = (unit, price)

Each time a *unit* (e.g., volume, file, record, page, kilobyte, image) is output, a cost of *price* is incurred.

Execute only

The user may execute a program but may not read, print, modify, or copy it. This rule protects against disclosure of an algorithm.

Enforcing an Authorized User List

In some cases, it is useful to have a rule that controls access to content for certain specific users or classes of users. For example, content may only be accessible to people over the age of eighteen, or to people having a rank greater than major, or to managers having a security clearance greater than TOP SECRET. In these cases, each user can be provided with a unique set of rules for that specific user. However, if the status of a user changes, then the rules for that user have to be changed. Accordingly, it is useful and convenient to have the rules be parameterized based on the status of the user and then have the user's status provided to the access mechanism in a secure fashion. This is a form of Role-Based Access Control.

Access Control Granularity

The above access control policies can be applied differently to various portions of the digital information. For example, a document's chapters might be controlled at different levels of quantity and quality; a map's information might be controlled differently at different latitudes and longitudes; portions of an image may be restricted in availability or resolution.

Controlling Distributions of Derivative Works

In many application environments where digital information is created, it is common to include extracts from other digital information. This is, for example, characteristic of the writing of scholarly papers, reviews, or regulations. The digital information containing the extract is a *derivative work*. The digital information from which the extract was copied is called the *parent work*.

PAC controls the distribution of derivative works. Creation of a derivative work can only be accomplished when permitted by the rules created by each of the owners of any data used in the derivative work. Use of a derivative work will, in general, require permissions from the owners of the derivative work as well as of the parent works.

The permissions associated with a work are incorporated into the permissions of any derivative work, either directly or by reference. License fees and restrictions imposed by the owner of a work are inherited by any derivative works. An n^{th} generation derivative work inherits the license fees and restrictions of each of its n-1 ancestors.

Controlling Use of Executable Software

PAC enables the creator of executable software to restrict the use of the software to only those who have acquired permissions for various software capabilities. Executable software is distributed in encrypted form, externally treated as data, as described above. In general, execution of a program can be controlled in a number of ways. Purchase of a license to execute software can be evidenced by a cryptographically protected certificate which is decrypted internally by the access mechanism. The executable software can check for the presence of the certificate, permission keys, or other information in the certificate once or many times during execution. Since the algorithm embodied in an executable program may be valuable digital information, the access mechanism can prevent a licensee from reading, copying, or modifying unencrypted executable code. To prevent disclosure of the unencrypted executable code, it is kept wholly within the security perimeter of the access mechanism for execution.

Control of Classified Data

PAC’s ability to support limitations on the primary and secondary distribution of data, access to data, and distribution of derivative data has obvious application to the protection of classified data. The payment feedback path can be augmented by an audit mechanism for tracking access. Similarly, the execution of classified programs, or programs operating on classified data, may be controlled by PAC.

Summary

Devising practical systems for controlled delivery of digital information from distributor to consumer has required innovation. This paper describes an invention, called Persistent Access Control, for controlling access to data, including derived data. Cryptography protects data during distribution and in storage. Access control rules are implemented in hardware. Tamper detection

prevents unauthorized access. PAC is applicable to national security and commercial environments, including electronic commerce.

Acknowledgements

The authors wish to express their appreciation to Bart Bailey and Dick Stewart who commented on previous drafts of this paper.

References

[1] Abrams, M., et al, "Generalized Framework for Access Control: Towards Prototyping the ORGCON Policy," *Proceedings National Computer Security Conference*, October 1991.

[2] Abrams, M. D., "Renewed Understanding of Access Control Policies," *Proceedings of the 16th National Computer Security Conference*, September 1993.

[3] Ciciora, W. S., "Inside the Set-Top Box," *IEEE Spectrum*, pp. 70-75, April 1995.

[4] Director of Central Intelligence, 4 May 1981, *Control of Dissemination of Intelligence Information*, Directive No. 1/7.

[5] Graubart, R., 1989, "On the Need for a Third Form of Access Control," *Proceedings of the 12th National Computer Security Conference*, pp. 296-303.

[6] McCollum, C. J., J. R. Messing, and L. Notargiacomo, 1990, "Beyond the Pale of MAC and DAC: Defining New Forms of Access Control," *Proceedings of the Symposium on Research in Security and Privacy*, IEEE Computer Society Press.

[7] Sandhu, R. S., 1992, "The Typed Access Matrix Model," *Proceedings of the Symposium on Research in Security and Privacy*, IEEE Computer Society, pp. 122-136.

[8] Sandhu, R. S., and G. S. Suri, 1992, "Implementation Considerations for the Typed Access Matrix Model in a Distributed Environment," *Proceedings of the 15th National Computer Security Conference*, pp. 221- 235

[9] Schneck, P. B., 1999, "Persistent Access Control to Prevent Piracy of Digital Information, *Proceedings of the IEEE*, Vol. 87, No. 7, July 1998, pp 1239-1250.

[10] Yee, B., *Using Secure Coprocessors*, Carnegie Mellon University, School of Computer Science, CMU-CS-94-149, 1994 (also available Defense Technical Information Center as AD-A281 255)