

the network either by implementing strict firewall rule sets or physically dividing the networks altogether. An additional short-term fix is to implement network-level encryption between routers that the ATM traffic traverses.

Long-term fixes involve changes in the application-level software. Protecting confidentiality requires encrypting all customer-related information that traverses the network. Ensuring data integrity requires better machine-to-machine authentication between the ATM and processor and the use of challenge-response protocols to counter replay attacks.

### 3.9 KEY TERMS, REVIEW QUESTIONS, AND PROBLEMS

#### Key Terms

✓ biometric	✓ identification	✓ smart card
✓ challenge-response protocol	✓ memory card	✓ static biometric
✓ claimant	✓ nonce	✓ subscriber
✓ credential	✓ password	✓ token
✓ credential service provider (CSP)	✓ rainbow table	✓ user authentication
✓ dynamic biometric	✓ registration authority (RA)	✓ verification
✓ enroll	✓ relying party (RP)	✓ verifier
✓ hashed password	✓ salt	
	✓ shadow password file	

#### Review Questions

- ✓ 3.1 In general terms, what are four means of authenticating a user's identity?
- ✓ 3.2 List and briefly describe the principal threats to the secrecy of passwords.
- ✓ 3.3 What is the significance of a shadow password file?
- ✓ 3.4 Explain how the proactive password checker approach can improve password security.
- ? 3.5 How can we classify the authentication protocols used with smart tokens?
- ✓ 3.6 List and briefly describe the principal physical characteristics used for biometric identification.
- ✓ 3.7 In the context of biometric user authentication, explain the terms, enrollment, verification, and identification.
- ✓ 3.8 How does remote user authentication differ from local authentication? Which one raised more security threats?
- ? 3.9 What is a Trojan horse attack? →

#### Problems

- 3.1 Explain the suitability or unsuitability of the following passwords:
  - a. qwerty    b. Einstein    c. wysiwyg (for "what you see is what you get")    d. drowssap
  - e. KVK 919    f. Florida    g. \*laptop\_admin#    h. cr@zyp@ss
- 3.2 An early attempt to force users to use less predictable passwords involved computer-supplied passwords. These passwords were generated using a pseudorandom

number generator. Suppose the passwords were nine-character long and were taken from the character set consisting of uppercase letters and digits so that the adversary has to search through all character strings of length 9 from a 36-character alphabet. Would a pseudorandom number generator with  $2^{16}$  possible starting values suffice? If yes, how? If not, then what should be the appropriate range for this pseudorandom number generator?

✓ 3.3 Assume that Personal Identification Numbers (PINs) are formed by nine-digit combinations of numbers 0 to 9. Assume that an adversary is able to attempt three PINs per second.

- Assuming no feedback to the adversary until each attempt has been completed, what is the expected time to discover the correct PIN?
- Assuming feedback to the adversary flagging an error as each incorrect digit is entered, what is the expected time to discover the correct PIN?

✗ 3.4 Assume source elements of length  $k$  are mapped in some uniform fashion into a target elements of length  $p$ . If each digit can take on one of  $r$  values, then the number of source elements is  $r^k$  and the number of target elements is the smaller number  $r^p$ . A particular source element  $x_i$  is mapped to a particular target element  $y_j$ .

- What is the probability that the correct source element can be selected by an adversary on one try?
- What is the probability that a different source element  $x_k$  ( $x_i \neq x_k$ ) that results in the same target element,  $y_j$ , could be produced by an adversary?
- What is the probability that the correct target element can be produced by an adversary on one try?

✗ 3.5 A phonetic password generator picks two segments randomly for each six-letter password. The form of each segment is CVC (consonant, vowel, consonant), where  $V = \langle a, e, i, o, u \rangle$  and  $C = \bar{V}$ .

- What is the total password population?
- What is the probability of an adversary guessing a password correctly?

✓ 3.6 Assume that credit card numbers are limited to the use of the 10 digits and that all numbers are 16 digits in length. Assume that an adversary needs around 31.69 years of time to test exhaustively all the possible credit card numbers. What is the rate at which the adversary is testing these numbers?

✓ 3.7 The NVIDIA Tesla K-20X GPU has 2688 cores, each operating at a 732-MHz frequency. Further, the GPU has 6 GB of DRAM with a bandwidth of 250 GB/sec that is shared among all the cores. If a password hashing scheme (PHS) takes 2 ms to compute a password:

- How many passwords can be tested by the GPU in one hour if the PHS consumes no memory?
- How many cores can work simultaneously if each hash computation requires 20 MB of DRAM? How many passwords can now be tested by the GPU in one hour?

✓ 3.8 The inclusion of the salt in the UNIX password scheme increases the difficulty of guessing by a factor of 4096. But the salt is stored in plaintext in the same entry as the corresponding ciphertext password. Therefore, those two characters are known to the attacker and need not be guessed. Why is it asserted that the salt increases security?

✓ 3.9 Assuming you have successfully answered the preceding problem and understand the significance of the salt, here is another question. Wouldn't it be possible to thwart completely all password crackers by dramatically increasing the salt size to, say, 24 or 48 bits?

✗ 3.10 Consider the Bloom filter discussed in Section 3.3. Define  $k$  = number of hash functions;  $N$  = number of bits in hash table; and  $D$  = number of words in dictionary.

- Show that the expected number of bits in the hash table that are equal to zero is expressed as

$$\phi = \left(1 - \frac{k}{N}\right)^D$$

- b. Show that the probability that an input word, not in the dictionary, will be falsely accepted as being in the dictionary is

$$P = (1 - \phi)^k$$

- c. Show that the preceding expression can be approximated as

$$P \approx (1 - e^{-kD/N})^k$$

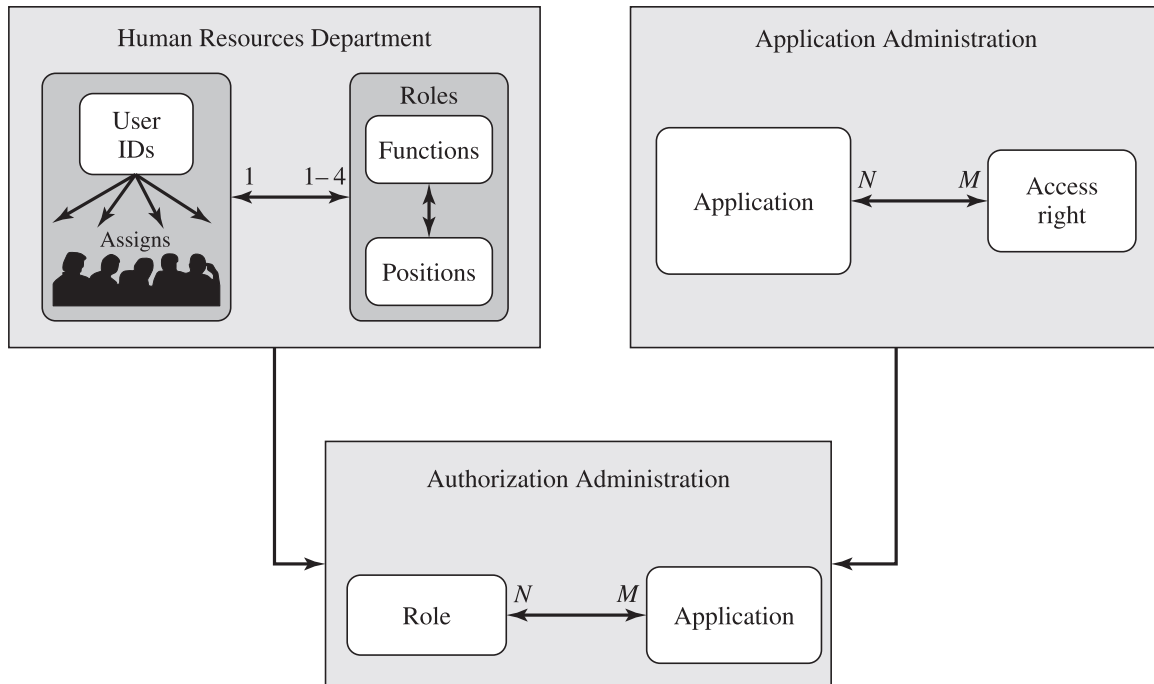
✓ 3.11

For the biometric authentication protocols illustrated in Figure 3.13, note the biometric capture device is authenticated in the case of a static biometric but not authenticated for a dynamic biometric. Explain why authentication is useful in the case of a stable biometric, but not needed in the case of a dynamic biometric.

How  
works

3.12

A relatively new authentication proposal is the Secure Quick Reliable Login (SQRL) described here: <https://www.grc.com/sqrl/sqrl.htm>. Write a brief summary of how SQRL works and indicate how it fits into the categories of types of user authentication listed in this chapter.



**Figure 4.14** Example of Access Control Administration

the applications the user invokes. This is illustrated in Table 4.5b. Role A has numerous access rights, but only a subset of those rights are applicable to each of the three applications that role A may invoke.

Some figures about this system are of interest. Within the bank, there are 65 official positions, ranging from a Clerk in a branch, through the Branch Manager, to a Member of the Board. These positions are combined with 368 different job functions provided by the human resources database. Potentially, there are 23,920 different roles, but the number of roles in current use is about 1,300. This is in line with the experience other RBAC implementations. On average, 42,000 security profiles are distributed to applications each day by the Authorization Administration module.

## 4.10 KEY TERMS, REVIEW QUESTIONS, AND PROBLEMS

### Key Terms

✓ access control	✓ attribute-based access control (ABAC)	✓ authorizations
✓ access control list	Attribute Exchange Network (AXN)	assessor ✗
✓ access management	✗	✓ capability ticket
✓ access matrix	✓ attribute provider	✗ cardinality
✓ access right	✓ auditor	✓ closed access control policy
✓ attribute		✓ credential

✓ credential management	✓ kernel mode	✓ prerequisite role
✓ discretionary access control (DAC)	✓ least privilege	✓ privilege
✗ dispute resolver	✓ <u>limited role hierarchy</u>	✓ protection domain
✓ <u>dynamic separation of duty (DSD)</u>	✓ mandatory access control (MAC)	✗ relying part
✗ entitlements	✓ mutually exclusive roles	✓ resource
✓ environment attribute	✓ object	✓ rights
✓ <u>general role hierarchy</u>	✓ object attribute	✓ role-based access control (RBAC)
✓ group	✗ open access control policy	✓ role constraints
✓ identity	Open Identity Exchange Corporation (OIX)	✓ role hierarchies
✗ <u>identity, credential, and access management (ICAM)</u>	Open Identity Trust Framework (OITF)	✓ separation of duty
✗ identity federation	OpenID	✓ session
✓ identity management	OpenID Foundation (OIDF)	✗ <u>static separation of duty (SSD)</u>
✓ identity provider	owner	✓ subject
✗ Information Card Foundation (ICF)	permission	✓ subject attribute
	policy	✗ <u>trust framework</u>
		✗ <u>trust framework provider</u>
		✓ user mode

### Review Questions

- ✓ 4.1 What is the difference between authentication and authorization?
- ✓ 4.2 How does RBAC relate to DAC and MAC?
- ✓ 4.3 List and define the three classes of subject in an access control system.
- ✓ 4.4 List and briefly explain the three basic elements of access control.
- ✓ 4.5 What is ABAC?
- ✗ 4.6 What is the difference between an access control list and a capability ticket?
- ✓ 4.7 List some of the main types of access control.
- ✓ 4.8 Briefly define the four RBAC models of Figure 4.8a.
- ✗ 4.9 What is meant by mutually exclusive roles in the RBAC<sub>3</sub> model?
- ✗ 4.10 Describe three types of role hierarchy constraints.
- ✗ 4.11 In the NIST RBAC model, what is the difference between SSD and DSD?

### Problems

- ✗ 4.1 For the DAC model discussed in Section 4.3, an alternative representation of the protection state is a directed graph. Each subject and each object in the protection state is represented by a node (a single node is used for an entity that is both subject and object). A directed line from a subject to an object indicates an access right, and the label on the link defines the access right.
- Draw a directed graph that corresponds to the access matrix of Figure 4.2a.
  - Draw a directed graph that corresponds to the access matrix of Figure 4.3.
  - Is there a one-to-one correspondence between the directed graph representation and the access matrix representation? Explain.

- 4.2 a. Explain, with an appropriate example, how protection domains provide flexibility.  
 b. How is the concept of protection domains related to operating systems? Explain by quoting an example from the UNIX operating system.

4.3 The VAX/VMS operating system makes use of four processor access modes to facilitate the protection and sharing of system resources among processes. The access mode determines:

- **Instruction execution privileges:** What instructions the processor may execute
- **Memory access privileges:** Which locations in virtual memory the current instruction may access

The four modes are as follows:

- **Kernel:** Executes the kernel of the VMS operating system, which includes memory management, interrupt handling, and I/O operations
- **Executive:** Executes many of the operating system service calls, including file and record (disk and tape) management routines
- **Supervisor:** Executes other operating system services, such as responses to user commands
- **User:** Executes user programs, plus utilities such as compilers, editors, linkers, and debuggers

A process executing in a less-privileged mode often needs to call a procedure that executes in a more-privileged mode; for example, a user program requires an operating system service. This call is achieved by using a change-mode (CHM) instruction, which causes an interrupt that transfers control to a routine at the new access mode. A return is made by executing the REI (return from exception or interrupt) instruction.

- a. A number of operating systems have two modes: kernel and user. What are the advantages and disadvantages of providing four modes instead of two?  
 b. Can you make a case for even more than four modes?

4.4 The VMS scheme discussed in the preceding problem is often referred to as a ring protection structure, as illustrated in Figure 4.15. Indeed, the simple kernel/user scheme is a two-ring structure. A disadvantage of a ring-structured access control system is that it violates the principle of least privilege. For example if we wish to have an object accessible in ring  $X$  but not ring  $Y$ , this requires that  $X < Y$ . Under this arrangement all objects accessible in ring  $X$  are also accessible in ring  $Y$ .

- a. Explain in more detail what the problem is and why least privilege is violated.  
 b. Suggest a way that a ring-structured operating system can deal with this problem.

4.5 UNIX treats file directories in the same fashion as files; that is, both are defined by the same type of data structure, called an inode. As with files, directories include a nine-bit protection string. If care is not taken, this can create access control problems. For example, consider a file with protection mode 644 (octal) contained in a directory with protection mode 730. How might the file be compromised in this case?

4.6 In the traditional UNIX file access model, which we describe in Section 4.4, UNIX systems provide a default setting for newly created files and directories, which the owner may later change. The default is typically full access for the owner combined with one of the following: no access for group and other, read/execute access for group and none for other, or read/execute access for both group and other. Briefly discuss the advantages and disadvantages of each of these cases, including an example of a type of organization where each would be appropriate.

4.7 Consider user accounts on a system with a Web server configured to provide access to user Web areas. In general, this uses a standard directory name, such as 'public\_html', in a user's home directory. This acts as their user Web area if it exists. However, to allow the Web server to access the pages in this directory, it must have at least search (execute) access to the user's home directory, read/execute access to the Web directory, and read access to any webpages in it. Consider the interaction of this requirement

UNIX  
passwords  
process  
threads

UNIX FS

Permissions  
of webserver  
directory structure



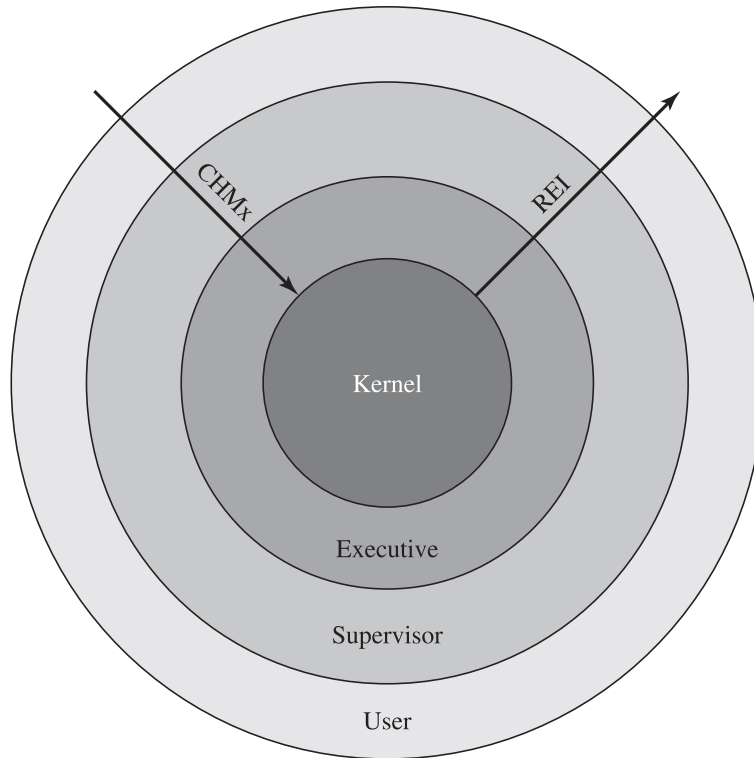


Figure 4.15 VAX/VMS Access Modes

with the cases you discussed for the preceding problem. What consequences does this requirement have? Note a Web server typically executes as a special user, and in a group that is not shared with most users on the system. Are there some circumstances when running such a Web service is simply not appropriate? Explain.

✓ 4.8

Assume an application requires access control policies based on the applicant's age and the type of funding to be provided. Using an ABAC approach, write policy rules for each of the following scenarios:

- If the applicant's age is more than 35, only "Research Grants (RG)" can be provided.
- If the applicant's age is less than or equal to 35, both "RG and Travel Grants (TG)" can be provided.

formal analysis of ABAC 4.9

Assume a system with  $K$  subject attributes,  $M$  object attributes and  $\text{Range}()$  denotes the range of possible values that each attribute can take. What are the number of roles and permissions required for an RBAC model? What is the problem with this approach if additional attributes are added?

✗ 4.10

For the NIST RBAC standard, we can define the general role hierarchy as follows:

$\text{RH} \subseteq \text{ROLES} \times \text{ROLES}$  is a partial order on ROLES called the inheritance relation, written as  $\geq$ , where  $r_1 \geq r_2$  only if all permissions of  $r_2$  are also permissions of  $r_1$ , and all users of  $r_1$  are also users of  $r_2$ . Define the set  $\text{authorized\_permissions}(r_i)$  to be the set of all permissions associated with role  $r_i$ . Define the set  $\text{authorized\_users}(r_i)$  to be the set of all users assigned to role  $r_i$ . Finally, node  $r_1$  is represented as an immediate descendant of  $r_2$  by  $r_1 \gg r_2$ , if  $r_1 \geq r_2$ , but no role in the role hierarchy lies between  $r_1$  and  $r_2$ .

- Using the preceding definitions, as needed, provide a formal definition of the general role hierarchy.
- Provide a formal definition of a limited role hierarchy.

~~4.11~~

In the example of Section 4.9, use the notation  $Role(x)$ ,  $Position$  and  $Role(x)$ ,  $Function$  to denote the position and the function associated with role  $x$ .

- a. We can define the role hierarchy for this example as one in which one role is superior to another if its position and functions are both superior. Express this relationship formally.
- b. An alternative role hierarchy is one in which a role is equal to another if its position is equal, regardless of the function. Express this relationship formally.

~~4.12~~

In the example of the online entertainment store in Section 4.6, with the finer-grained policy that includes premium and regular users, describe the ABAC policy rules for accessing a movie, and list all the advantages of an ABAC control policy.



## 5.9 KEY TERMS, REVIEW QUESTIONS, AND PROBLEMS

### Key Terms

✓ attribute	✓ inband attack	✓ run-time prevention
✓ blind SQL injection	✓ inference	✓ Structured Query Language (SQL)
✓ cascading authorizations	✓ inference channel	✓ SQL injection (SQLi) attack
✓ compromise	✓ inferential attack	✓ tautology
<del>data center</del>	✓ out-of-band attack	✓ tuple ( — , — , — )
✓ data swapping	✓ parameterized query insertion	✓ view
✓ database	<del>partitioning</del>	
✓ database access control	✓ piggybacked queries	
✓ database encryption	✓ primary key	
✓ database management system (DBMS)	✓ query language	
✓ defensive coding	✓ query set	
✓ detection	✓ relation	
✓ end-of-line comment	✓ relational database	
✓ foreign key	✓ relational database management system (RDBMS)	

### Review Questions

- ✓ 5.1 Define the terms *database*, *database management system*, and *query language*.
- ✓ 5.2 What is a relational database and what are its principal ingredients?
- ✓ 5.3 What is an SQL injection attack?
- ✓ 5.4 What are the implications of an SQL injection attack?
- ✓ 5.5 List the categories for grouping different types of SQLi attacks.
- ✓ 5.6 Why is RBAC considered fit for database access control?
- ✓ 5.7 State the different levels at which encryption can be applied to a database.
- ~~5.8~~ List and briefly define four data center availability tiers.

*cluster of clusters*

### Problems

- ✓ 5.1 Consider a simplified database for an organization that includes information of several departments (identity, name, manager, number of employees) and of managers and employees of the respective departments. Suggest a relational database for efficiently managing this information.
- 5.2 The following table provides information on students of a computer programming club.

Student-ID	Name	Skill Level	Age
99	Jimmy	Beginner	20
36	David	Experienced	22
82	Oliver	Medium	21
23	Alice	Experienced	21

✓ The primary key is *Student-ID*. Explain whether or not each of the following rows can be added to the table.

Student-ID	Name	Skill Level	Age
91	Tom	Experienced	22
36	Dave	Experienced	21
✓	Bob	Beginner	20

5.3 The following table shows a list of cars and their owners that is used by a car service station.

C_Name	Model	Company	DOP	Owner	O_Phone	O_E-mail
Camaro	2LS	Chevrolet	9/9/06	David	2132133	dd@abc.com
Falcon	XR6	Ford	2/21/07	Dave	1245513	dv@abc.com
Cruze	LT	Chevrolet	5/12/12	David	1452321	dd@abc.com
Camaro	2LT	Chevrolet	7/6/10	Alice	3253254	al@ab.com
Roadster	Roadster	Tesla	1/20/13	Dave	2353253	dv@abc.com
Focus	S	Ford	4/10/12	Oliver	3251666	ol@abc.com
Model X	Model X	Tesla	3/9/14	Bob	7567443	bb@abc.com

✓ a. Describe the problems that are likely to occur when using this table.

✓ b. Break the table into two tables in a way that fixes the problems.

✓ 5.4 We wish to create an employee table containing the employee's ID number, first name, last name, and department. Write an SQL statement to accomplish this.

✓ 5.5 Consider an SQL statement:

SELECT id, forename, surname FROM authors WHERE forename = 'david' AND id = 939

✓ a. What is this statement trying to search from the database?

✓ b. Assume that the firstname and id fields are being gathered from user-supplied input, and suppose the user responds with:  
Firstname: david'; drop table employees - -  
id: 939;

What will be the effect?

✓ c. Now suppose the user responds with:

firstname: ' OR 9 = 9 - -

id: 939

What will be the effect?

✓ 5.6 Figure 5.14 shows a fragment of code that implements the login functionality for a database application. The code dynamically builds an SQL query and submits it to a database.

✓ a. Suppose a user submits login, password, and pin as Mike, Mike@256, and 4242. Write the SQL query that is generated.

✓ b. If, instead of the previous inputs, the user submits for each of the login, password and pin fields:

' or '' = '

What is the effect?

```

1. String login, password, pin, query
2. login = getParameter("login");
3. password = getParameter("pass");
3. pin = getParameter("pin");
4. Connection conn.createConnection("MyDataBase");
5. query = "SELECT accounts FROM users WHERE login='" +
6.     login + "' AND pass = '" + password +
7.     "' AND pin=" + pin;
8. ResultSet result = conn.executeQuery(query);
9. if (result!=NULL)
10     displayAccounts(result);
11 else
12     displayAuthFailed();

```

*See Problem 56*

**Figure 5.14** Code for Generating an SQL Query

✓ 5.7 The EXISTS operator is used to test for the existence of any record in a subquery. Suppose you know that a user with the login Mike exists in the user table but you do not know their password. You enter the following in the login field:

' OR EXISTS (SELECT \* FROM users WHERE name = 'Mike' AND password LIKE '%t%') –

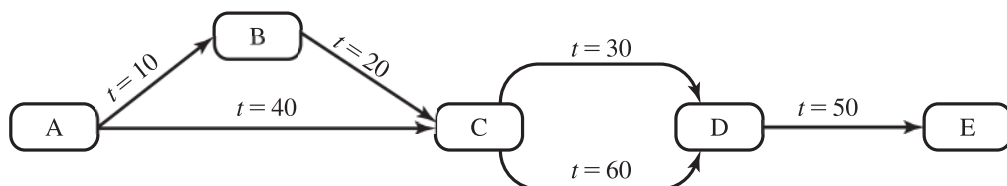
What is the effect?

✓ 5.8 Assume A, B, and C grant certain privileges on the employee table to X, who in turn grants them to Y, as shown in the following table, with the numerical entries indicating the time of granting:

UserID	Table	Grantor	READ	INSERT	DELETE
X	Employee	A	15	15	—
X	Employee	B	20	—	20
Y	Employee	X	25	25	25
X	Employee	C	30	—	30

At time  $t = 35$ , B issues the command REVOKE ALL RIGHTS ON Employee FROM X. Which access rights, if any, of Y must be revoked, using the conventions defined in Section 5.2?

✓ 5.9 Figure 5.15 shows a sequence of grant operations for a specific access right on a table. Assume at  $t = 70$ , B revokes the access right from C. Using the conventions defined in Section 5.2, show the resulting diagram of access right dependencies.



**Figure 5.15** Cascaded Privileges

- ✓ 5.10 Figure 5.16 shows an alternative convention for handling revocations of the type illustrated in Figure 5.6.

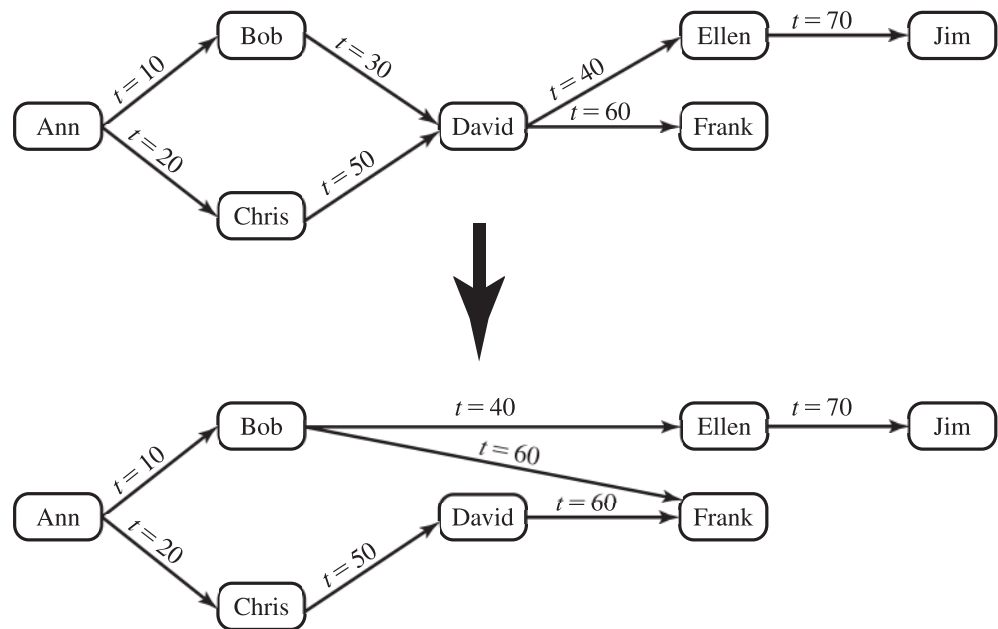


Figure 5.16 Bob Revokes Privilege from David, Second Version

- Steps
- a. Describe an algorithm for revocation that fits this figure.
- b. Compare the relative advantages and disadvantages of this method to the original method, illustrated in Figure 5.6.
- ✓ 5.11 Consider the parts department of a plumbing contractor. The department maintains an inventory database that includes parts information (part number, description, color, size, number in stock, etc.) and information on vendors from whom parts are obtained (name, address, pending purchase orders, closed purchase orders, etc.). In an RBAC system, suppose roles are defined for accounts payable clerk, an installation foreman, and a receiving clerk. For each role, indicate which items should be accessible for read-only and read-write access.
- 5.12 Imagine you are the database administrator for a military transportation system. You have a table named cargo in your database that contains information on the various cargo holds available on each outbound airplane. Each row in the table represents a single shipment and lists the contents of that shipment and the flight identification number. Only one shipment per hold is allowed. The flight identification number may be cross-referenced with other tables to determine the origin, destination, flight time, and similar data. The cargo table appears as follows:

Flight ID	Cargo Hold	Contents	Classification
1254	A	Boots	Unclassified
1254	B	Guns	Unclassified
1254	C	Atomic bomb	Top Secret
1254	D	Butter	Unclassified

Inference?

Suppose two roles are defined: Role 1 has full access rights to the cargo table. Role 2 has full access rights only to rows of the table in which the Classification field has the value Unclassified. Describe a scenario in which a user assigned to role 2 uses one or more queries to determine that there is a classified shipment on board the aircraft.

✓ 5.13

Users hulkhogan and undertaker do not have the SELECT access right to the Inventory table and the Item table. These tables were created by and are owned by user bruno-s. Write the SQL commands that would enable bruno-s to grant SELECT access to these tables to hulkhogan and undertaker.

✓ 5.14

In the example of Section 5.6 involving the addition of a start-date column to a set of tables defining employee information, it was stated that a straightforward way to remove the inference channel is to add the start-date column to the employees table. Suggest another way.

✓ 5.15

Consider a database table that includes a salary attribute. Suppose the three queries **sum**, **count**, and **max** (in that order) are made on the salary attribute, all conditioned on the same predicate involving other attributes. That is, a specific subset of records is selected and the three queries are performed on that subset. Suppose the first two queries are answered, and the third query is denied. Is any information leaked?

See  
Handouts  
on  
Inference

filtering software of a LAN router or switch. As with ingress monitors, the external firewall or a honeypot can house the monitoring software. Indeed, the two types of monitors can be installed in one device. The egress monitor is designed to catch the source of a malware attack by monitoring outgoing traffic for signs of scanning or other suspicious behavior. This monitoring could look for the common sequential or random scanning behavior used by worms and rate limit or block it. It may also be able to detect and respond to abnormally high e-mail traffic such as that used by mass e-mail worms, or spam payloads. It may also implement data exfiltration “data-loss” technical counter measures, monitoring for unauthorized transmission of sensitive information out of the organization.

Perimeter monitoring can also assist in detecting and responding to botnet activity by detecting abnormal traffic patterns associated with this activity. Once bots are activated and an attack is underway, such monitoring can be used to detect the attack. However, the primary objective is to try to detect and disable the botnet during its construction phase, using the various scanning techniques we have just discussed, identifying and blocking the malware that is used to propagate this type of payload.

### Distributed Intelligence Gathering Approaches

The final location where anti-virus software is used is in a distributed configuration. It gathers data from a large number of both host-based and perimeter sensors, relays this intelligence to a central analysis system able to correlate and analyze the data, which can then return updated signatures and behavior patterns to enable all of the coordinated systems to respond and defend against malware attacks. A number of such systems have been proposed. This is a specific example of a distributed intrusion prevention system (IPS), targeting malware, which we will discuss further in Section 9.6.

## 6.11 KEY TERMS, REVIEW QUESTIONS, AND PROBLEMS

### Key Terms

advanced persistent threat	infection vector	ransomware
adware	keyloggers	rootkit
attack kit	logic bomb	scanning
backdoor	macro virus	spear-phishing
blended attack	malicious software	spyware
boot-sector infector	malware	stealth virus
bot	metamorphic virus	trapdoor
botnet	mobile code	Trojan horse
crimeware	parasitic virus	virus
data exfiltration	payload	watering-hole attack
downloader	phishing	worm
drive-by-download	polymorphic virus	zombie
e-mail virus	propagate	zero-day exploit



## Review Questions

- 6.1 What are three broad mechanisms that malware can use to propagate?
- 6.2 What are four broad categories of payloads that malware may carry?
- 6.3 What characteristics of an advanced persistent threat give it that name?
- 6.4 What are typical phases of operation of a virus or worm?
- 6.5 What is a blended attack?
- 6.6 What is the difference between a worm and a zombie?
- 6.7 What does “fingerprinting” mean for network worms?
- 6.8 What is a “drive-by-download” and how does it differ from a worm?
- 6.9 How does a Trojan enable malware to propagate? How common are Trojans on computer systems? Or on mobile platforms?
- 6.10 What is a “logic bomb”?
- 6.11 What is the difference between a backdoor, a bot, a keylogger, spyware, and a rootkit? Can they all be present in the same malware?
- 6.12 What is the difference between a “phishing” attack and a “spear-phishing” attack, particularly in terms of who the target may be?
- 6.13 What is a clickjacking vulnerability?
- 6.14 List a few characteristics to classify rootkits.
- 6.15 Briefly describe the elements of a GD scanner.
- 6.16 Describe some rootkit countermeasures.

## Problems

- 6.1 A computer virus places a copy of itself into other programs, and arranges for that code to be run when the program executes. The “simple” approach just appends the code after the existing code, and changes the address where code execution starts. This will clearly increase the size of the program, which is easily observed. Investigate and briefly list some other approaches that do not change the size of the program.
- 6.2 The question arises as to whether it is possible to develop a program that can analyze a piece of software to determine if it is a virus. Consider that we have a program D that is supposed to be able to do that. That is, for any program P, if we run D(P), the result returned is TRUE (P is a virus) or FALSE (P is not a virus). Now consider the following program:

```

Program CV :=
{. . .
  main-program :=
    {if D(CV) then goto next:
      else infect-executable;
    }
  next:
}
```

In the preceding program, infect-executable is a module that scans memory for executable programs and replicates itself in those programs. Determine if D can correctly decide whether CV is a virus.

- 6.3 The following code fragments show a sequence of virus instructions and a metamorphic version of the virus. Describe the effect produced by the metamorphic code.

Original Code	Metamorphic Code
<pre> mov eax, 5 add eax, ebx call [eax]</pre>	<pre> mov eax, 5 push ecx pop ecx add eax, ebx swap eax, ebx swap ebx, eax call [eax] nop</pre>

- 6.4 The list of passwords used by the Morris worm is provided at this book's website.
- The assumption has been expressed by many people that this list represents words commonly used as passwords. Does this seem likely? Justify your answer.
  - If the list does not reflect commonly used passwords, suggest some approaches that Morris may have used to construct the list.
- 6.5 Consider the following fragment:

```

legitimate code
if an infected document is opened;
    trigger_code_to_infect_other_documents();
legitimate code
```

What type of malware is this?

- 6.6 Consider the following fragment embedded in a webpage:

```

username = read_username();
password = read_password();
if username and password are valid
    return ALLOW_LOGIN;
    executable_start_download();
else return DENY_LOGIN
    executable_start_download();
```

What type of malicious software is this?

- 6.7 Many websites use a CAPTCHA image on their login page. A typical application of this is in an HTML form asking for the email ID and the login password of a user. The webpage also shows some numbers and letters, modified in a manner such that it is still easy for a human to recognize these characters. The user is then asked to recognize these characters and is granted login access only when they successfully enter the characters. Explain how using a CAPTCHA can help prevent email spam. What is the main difficulty with using CAPTCHAs?
- 6.8 What are honeypots? How are they better at resisting spam bots than CAPTCHAs?
- 6.9 Suppose that while working on a course assignment you come across a software that seems efficient to complete the assignment. When you run the software, however, you observe it keeps redirecting you to a different website and does not do the desired task. Is there a threat to your computer system?

- 6.10** Suppose you have a new smartphone and are excited about the range of apps available for it. You read about a really interesting new game that is available for your phone. You do a quick Web search for it and see that a version is available from one of the free marketplaces. When you download and start to install this app, you are asked to approve the access permissions granted to it. You see that it wants permission to “Send SMS messages” and to “Access your address-book.” Should you be suspicious that a game wants these types of permissions? What threat might the app pose to your smartphone, should you grant these permissions and proceed to install it? What types of malware might it be?
- 6.11** Assume you receive an e-mail, which appears to come from a senior manager in your company, with a subject indicating that it concerns a project that you are currently working on. When you view the e-mail, you see that it asks you to review the attached revised press release, supplied as a PDF document, to check that all details are correct before management releases it. When you attempt to open the PDF, the viewer pops up a dialog labeled “Launch File” indicating that “the file and its viewer application are set to be launched by this PDF file.” In the section of this dialog labeled “File,” there are a number of blank lines, and finally the text “Click the ‘Open’ button to view this document.” You also note that there is a vertical scroll-bar visible for this region. What type of threat might this pose to your computer system should you indeed select the “Open” button? How could you check your suspicions without threatening your system? What type of attack is this type of message associated with? How many people are likely to have received this particular e-mail?
- 6.12** Assume you receive an e-mail, which appears to come from an online air ticket reservation system, includes original logo and has following contents: “Dear Customer, Thank you for booking your air ticket through our online reservation system. The PNR for your journey from City1 to City2 is JADSA and for your return journey is EWTEQ. You can download your tickets by logging in through this *link*.” Assume you are a frequent visitor of City1 and City2 is another city you visit very frequently. What form of attack is this e-mail attempting? What is the most likely mechanism used to distribute this e-mail? How should you respond to such e-mails?
- 6.13** Suppose you receive a letter, which appears to come from your company’s mail server stating that the password for your account has been changed, and that an action is required to confirm this. However, as far as you know, you have not changed the password! What may have occurred that led to the password being changed? What type of malware, and on which computer systems, might have provided the necessary information to an attacker that enabled them to successfully change the password?
- 6.14** One of the possible locations to deploy anti-virus software is an organization’s firewall so that it can obtain a larger view of the malware activity. Describe at least one limitation of adopting this approach of deploying the anti-virus software. What are the possible ways, if any, to overcome this limitation?