AP Computer Science Principles

Summer Assignment

2023-2024 School Year

Welcome to Advanced Placement Computer Science Principles! In this course, we will discuss some of the central ideas of computer science and computational thinking, including abstraction, algorithmic thinking, data, the Internet, and the global impact of the computer revolution.

To prepare for the course, you must complete this assignment over the summer. This summer assignment is *due before 3:00pm on the **first B-day of school** on Schoology for both A-day and B-day students* and will form the basis of class discussion and activity during Unit 1 and Unit 6, in which we will discuss computational thinking and logical reasoning, describe computational processes using algorithms and 'pseudocode', and investigate encryption and cybersecurity.

In Part 1 you will read about, learn, analyze and practice using some very simple encryption algorithms by hand. For Part 2 of the project you will study and investigate cybersecurity.

Summer Assignment Rubric

	Points	**Points Earned**
Part I — Encryption	**50**	
Worksheet #1	10	
Worksheet #2	10	
Make Your Own Cipher	30	
Part II — Cybersecurity	**50**	
Reflection #1	2	
CIA Components Multiple Choice Questions	3	
Reflection #2	2	
Reflection #3	2	
Investigate	41	
Total	**100**	

Part 1A – Encryption

The Need for Secure Communication

Encryption is the process of encoding data to prevent unauthorized access.

Decryption is the process of decoding the data.

For more than three millennia, humans have sought ways to protect and secure information so that it is only available to select individuals. From ancient military leaders seeking to keep their battle plans out of enemy hands or modern-day financial investors trading stocks online, people have needed to restrict access to information.

One way to ensure that information is not accessible is to completely destroy the information. Imagine writing a private message down on a sheet of paper and then shredding the page into a million tiny pieces. The information contained within the original message will be all but lost, ensuring that no bystander might piece it all back together again. Unfortunately, it means that you will also lose your data, rendering this an undesirable solution to the problem.

Instead, the real problem is to obscure the data without destroying it. If only there were a way to scramble the information in some way that renders it unreadable to anyone who does not know how to unscramble it. This is the goal of cryptology, the study of securing, or *encrypting,* information so that it is inaccessible by third parties. Encryption is the process of encoding data to prevent unauthorized access. Decryption is the process of decoding the data.

Alice, Bob, and Eve

The classic example that is often used to demonstrate and explain the problem of encryption is the case of Alice, Bob, and Eve. In this scenario, Alice wishes to send a message to Bob, but she and Bob are unable to meet in person to exchange information directly. Instead, all communications between Alice and Bob must go through a third-party messenger, Eve. Unfortunately, Eve (as in eavesdropper) is potentially untrustworthy and might read any messages she is asked to deliver.

This hypothetical scenario is a model of what happens every time you make a private phone call, access a web page, or send an email or text message. You, like Alice, are reliant upon the computers, routers, online services, and other parts of the network infrastructure to deliver the data of your message to your intended recipient. Unfortunately, as with Eve, you have no oversight of these components and cannot guarantee that an interested third party is not accessing the private data you are transmitting.

Caesar Cipher

One of the earliest and simplest attempts at encryption is the Caesar cipher, employed by Julius Caesar in the 1st century BCE. This encryption method is known as a substitution cipher because it substitutes each letter of the original, unencrypted message, the *plaintext,* with a corresponding letter in the final, encrypted message, the *ciphertext*.

The Caesar cipher works by aligning two alphabets against one another and offsetting them by a number of positions. Caesar himself used a "left rotation" of three spaces, causing an `a` of the plaintext to align with an `x` in the ciphertext.

```
Plaintext:   abcdefghijklmnopqrstuvwxyz

Ciphertext:   xyzabcdefghijklmnopqrstuvw
```

Try playing around with an interactive demo of the Caesar cipher (Links to an external site.)

to see how messages can be encrypted and decrypted.

For example, if you enter a plaintext message of `this is a caesar cipher` and set the offset to be 23 (left shift of 3), then clicking on "Encipher Plaintext," will produce the following ciphertext:

```
Plaintext: this is a caesar cipher

Ciphertext: QEFP FP X ZXBPXO ZFMEBO
```

Similarly, entering a ciphertext of `GUVF VF QRPELCGRQ` and an offset of 13 will produce the following plaintext:

```
Ciphertext: GUVF VF QRPELCGRQ

 Plaintext: this is decrypted
```

Keys

Notice that in order to decrypt each of these messages, you must know three things:

1. The ciphertext message
2. The method of encryption used to create the ciphertext (such as a Caesar cipher)
3. The number of positions by which the plaintext and ciphertext alphabets have been offset

The last of these items, the offset, serves as the *key* that effectively locks and unlocks the message. Symmetric key encryption involves one key for both encryption and decryption.

Using the key to encrypt the message into a ciphertext secures the message and protects it from prying eyes, much like locking the message in a box or safe. Likewise, you use the key to properly align the alphabets and unlock the message in order to read the original plaintext. Both parties need to know the key for symmetric encryption.

In the earlier scenario, as long as Bob knows which key Alice used to encrypt her message, he can use the same key to decrypt the message once he receives it. However, without being told which key was used, Eve cannot decrypt the message as easily as either Bob or Alice.

Obviously, as you have previously seen for yourself, a Caesar cipher's key can be easily deduced through simple trial and error. After all, there are only 25 possible keys and it is easy to try each one through brute force.

Today, modern encryption schemes are far more sophisticated than those used in Julius Caesar's day, and they use far more complex keys that are much harder to guess through brute force techniques. But the model used more than 2000 years ago is still more or less the one we use today.

1. A sender (Alice) uses an encryption scheme and a key to encode a message.
2. The encoded message is transmitted through one or more intermediate and potentially untrustworthy handlers (Eve).
3. The recipient, who knows which encryption scheme was used and is already in possession of the necessary key, decrypts the message back into its original, plaintext form.

The only difference between the ancient Roman times and now is that computational processing power has made Eve's job much easier and the need for stronger encryption algorithms and keys much greater.

A little more about Caesar Ciphers

A Caesar cipher is a replacement cipher that replaces each letter in the plaintext message with a different one. Typically, to create a key for a Caesar cipher, you create two rows of letters. The first row is just the alphabet, from A to Z. The second row is also the alphabet, but shifted by a number of characters. The example below shows a Caesar cipher with shift +6:

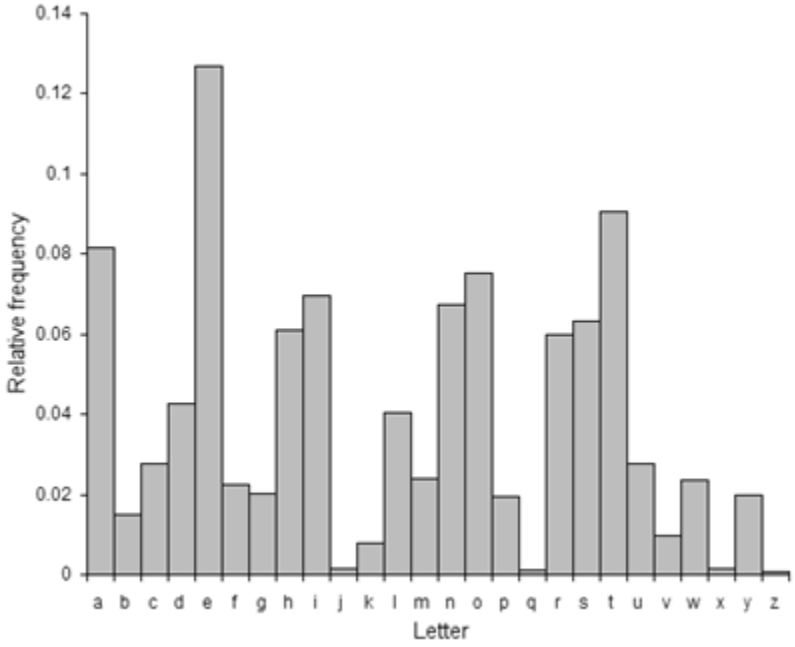
ABCDEFGHIJKLMNOPQRSTUVWXYZ

GHIJKLMNOPQRSTUVWXYZABCDEF

In this case, each letter A in the plaintext message would be represented by a G in the ciphertext, a "shift' of 6 letters forward in the alphabet. Each B would be represented by an H, each C by an I, and so on.

You can create different Caesar ciphers by shifting the second row different amounts. For example, a shift of +12 would encode A with M, B with N, and so on.

Caesar ciphers can be easily cracked if you know the shift used to create the ciphertext. Even if you don't, you can use frequency analysis. The chart below shows the relative frequency of each letter's appearance in normal English text. If you don't know the shift amount, you could try to find the most common letters in the ciphertext and assume they represent the most common letters in the plaintext English – E or T.



Vigenère Cipher

A Caesar cipher may be sufficient for obscuring notes passed in class, but modern technology makes circumventing such simple algorithms quite easy. After all, there are 25 possible offsets to attempt in a simple Caesar cipher based on the English alphabet. More sophisticated algorithms were developed centuries ago, such as the Vigenère cipher, (Links to an external site.) which effectively rotates the Caesar cipher offset used to encrypt each new letter in a text. However, encryption is just a small piece of the puzzle to secure communication. Cybersecurity (Links to an external site.) —"measures taken to protect a computer or computer system (as on the internet) against unauthorized access or attack"—encompasses a much broader scope of techniques.

Traditionally, the CIA Triad (Links to an external site.) defines the target areas when developing a secure system. CIA is an initialism representing three concepts:

- Confidentiality is the ability to limit access to information to a certain set of users.

- Integrity is the certainty that information is accurate.
- Availability is the reliability of access to information.

Although the field of information security has grown to include other areas of focus, such as authentication (Links to an external site.), these three core concepts remain central.

Confidentiality

Confidentiality is the ability to limit access to information to a certain set of users.

Confidentiality is what most people equate with cybersecurity, and at the heart of confidentiality protocols lie encryption. We examined encryption through the Caesar cipher as an example application of an algorithm. Although the basic algorithm of the Caesar cipher, using mathematics to convert one character to another, remains relevant today, the protocols and mathematics used to apply the algorithm have gotten much more sophisticated.

Keys

The first departure from the Caesar cipher algorithm is the use of more sophisticated *keys* in generating an encrypted text. The Vigenère cipher (Links to an external site.) took the basic idea of a Caesar cipher and added an extra piece of information to make it more secure—a keyphrase.

In this simple example comparing the two methods, imagine that a plaintext message, `IT IS BURIED IN THE BACKYARD` is encrypted first with a Caesar cipher with an offset of 3:

```
Plaintext:  IT IS BURIED IN THE BACKYARD
```

```
Cyphertext:  LW LV EXULHG LQ WKH EDFNBDUG
```

and second with the Vigenère cipher using the keyphrase `DIG`:

```
Plaintext:  IT IS BURIED IN THE BACKYARD
```

```
Cyphertext:  LB OV JAUQKG QT WPK EIINGGUL
```

In reality, this Vigenère cipher is just using *three different Caesar ciphers* in succession, each with an offset corresponding to the letters of the keyphrase, `DIG`:

```
'D' is 3 letters after the beginning of the alphabet, 'A', so offset by 3
```

```
'I' is 8 letters after the beginning of the alphabet, 'A', so offset by 8
```

```
'G' is 6 letters after the beginning of the alphabet, 'A', so offset by 6
```

Then, the cycle repeats:

```
Plaintext:  IT IS BURIED IN THE BACKYARD

Keyphrase:  DI GD IGDIGD IG DIG DIGDIGDI

Offset:         38 63 863863 86 386 38638638

Cyphertext: LB OV JAUQKG QT WPK EIINGGUL
```

It is worth noting that the Vigenère cipher with the keyphrase of simply the letter `D` is the same as our original Caesar cipher in this example—each time offsetting by 3.

Restricted Knowledge

Notice that each of these encryption methods relies on restricted knowledge.

With the Caesar cipher, anyone who knows the algorithm and the offset can decrypt the message. However, we have seen that even if you know just the algorithm, it is not hard to decrypt without also knowing the offset—there are only 25 possibilities. With the Vigenère cipher, anyone who knows the algorithm and the key can decrypt the message. It is much more difficult to decrypt the message without the key than before because patterns in the text are less obvious, and there are many more possibilities than 25.

Complete Worksheet #1 at the end of the packet

Matrix Transposition Ciphers

Matrix Transposition ciphers involve writing the plaintext message into the rows of a matrix, then reading down the columns of the matrix to obtain the ciphertext. In this way, the same letters of the plaintext message are present in the ciphertext, but they are rearranged into a different order. In the example below the message COMPUTER SCIENCE IS COOL is entered into the matrix in rows of 4:

C	O	M	P
U	T	E	R
S	C	I	E
N	C	E	I
S	C	O	O
L	X	X	X

Notice that the message did not fill the matrix, so the 'null character' X was used to complete the last row. The ciphertext is now created by reading down the columns:

CUSNS LOTCC CXMEI EOXPR EIOX

The ciphertext is written in groups of five letters here for ease of reading, regardless of the length of the original plaintext words or the size of the matrix used. To decode, we would need to enter the ciphertext back into the columns a matrix of the appropriate size, then read across the rows.

Revised 04/2023

Matrix Transposition ciphers can be made more complicated by mixing up the order the columns are read to create the ciphertext, usually by using a 'keyword'. For example, I could encode the message EVERYONE LOVES ICE CREAM using the keyword SUNDAE.

The keyword is written across the top, and the plaintext entered into rows below. To create the ciphertext, we read down the columns, but the order this time is determined by the keyword SUNDAE. We will go alphabetically through the keyword, starting with the column under the A, then the D, then E, and so on:

YBCXR OEXOE RXELC MENSE VEIA

To decode the message, the receiver would need to know the keyword, so that they can use the correct size matrix rows, and so that they know the order in which to enter the ciphertext into the columns.

Complete Worksheets #2-3 at the end of the packet

Part 2– Cybersecurity

Restricted Information

Encryption

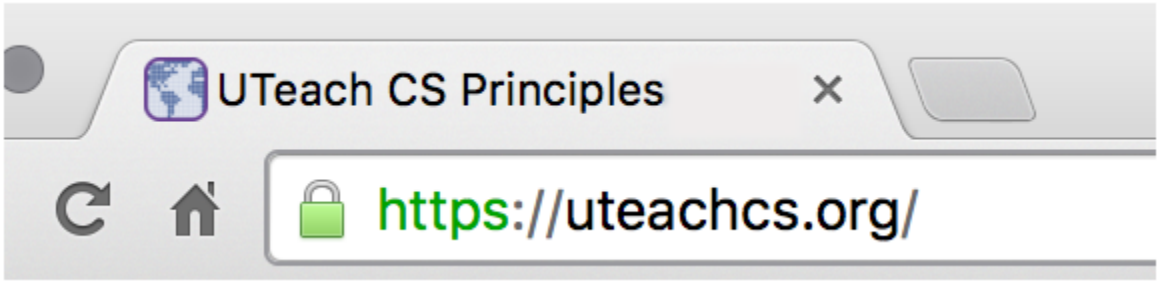
As we learned back in Unit 1, encryption is one security measure used to protect our digital information. There are two types of encryptions used to send secure messages: symmetric key encryption and asymmetric key encryption.

Watch the video to learn more: https://www.youtube.com/watch?v=AQDCe585Lnc&t=2s

As seen in the video, symmetric key encryption has a major flaw: there is one key needed for both encryption and decryption. Thus, how to securely get this key to the recipient becomes quite the quandary.

> **Public-key encryption** (asymmetric) pairs a public key for encryption and a private key for decryption. The sender does not need the receiver's private key to encrypt a message, but the receiver's private key is required to decrypt the message. This provides much more secure communication.

Modern secure message transmission through the internet uses an asymmetric key system. Certificate authorities (CAs) issue digital certificates that validate the ownership of encrypted keys used in secured communications and are based on a trust model. When a website uses this type of security protocol, you will see a green padlock icon in the browser's address bar. This is called Secure Sockets Layer (SSL). When you visit websites with the padlock (sometimes green or gray), you can be confident the web host is following the protocol for security. If you do not see the padlock, enter at your own risk!



Complete Reflection #1 at the end of the packet

CIA Triad

Encryption is just a small piece of the puzzle to secure communication. **Cybersecurity**—measures taken to protect a computer or computer system (as on the internet) against unauthorized access or attack—encompasses a much broader scope of techniques.

Traditionally, the **CIA Triad** defines the target areas when developing a secure system. CIA is an initialism representing the following concepts:

- **C**onfidentiality is the ability to limit access to information to a certain set of users.
- **I**ntegrity is the certainty that information is accurate.
- **A**vailability is the reliability of access to information.

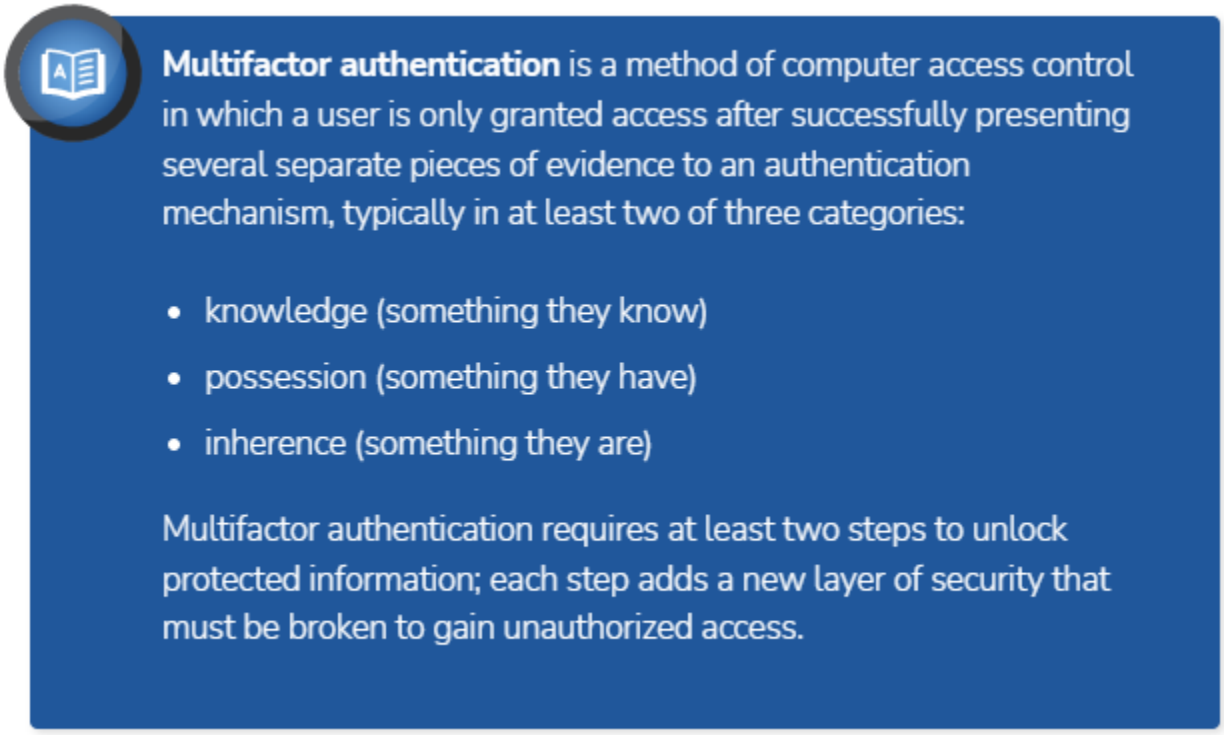
Watch this video to learn more: https://www.youtube.com/watch?v=rwigKjEsdTc

The field of information security has grown to include other areas of focus, such as **authentication.** These three core concepts along with authentication remain central to cybersecurity.

Complete CIA Components Multiple Choice Questions at the end of the packet

Authentication

Authentication measures protect devices and information from unauthorized access. Examples of authentication measures include strong passwords and *multifactor authentication*. As seen in the first unit project, a strong password can be created using an algorithm instead of using the same password for every site.

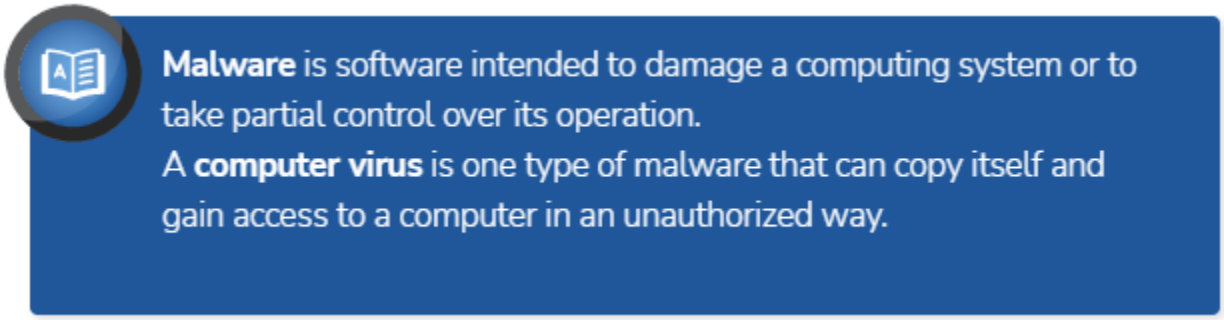


Watch this video to learn more: https://www.youtube.com/watch?v=tFv101qURKE

Complete Reflection #2 at the end of the packet

Other Security Risks

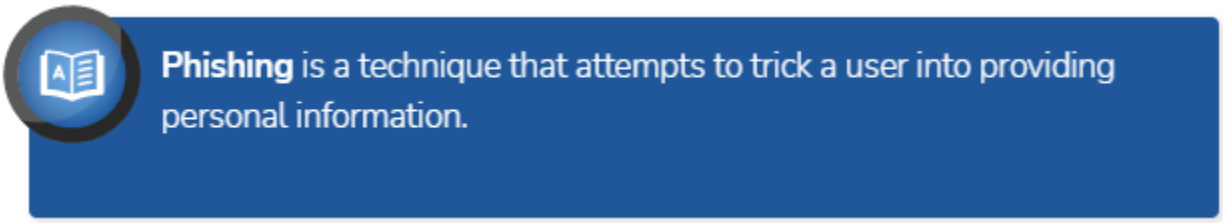
Malware and Viruses



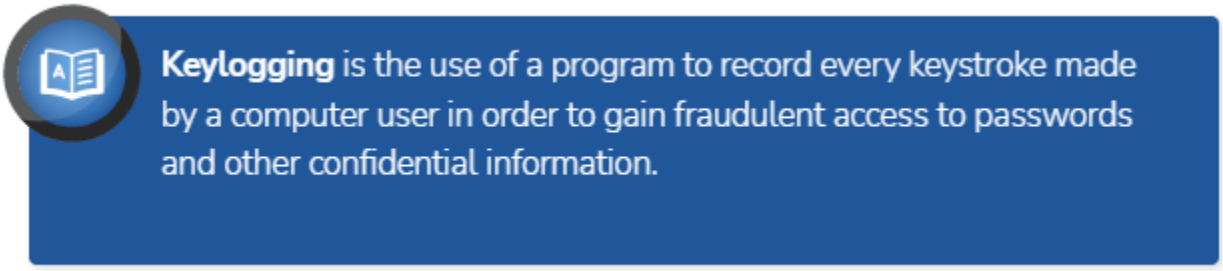
Computer viruses often attach themselves to legitimate programs and start running independently on a computer. Generic protection against these kinds of attacks exists in the form of software that scans your computer and prevents these infections. Some common ways to get one of these infections is through untrustworthy downloads from freeware or shareware sites or from a malicious link on a familiar-looking website or email message. Unsolicited emails or attachments, links, and forms in emails can be used to compromise the security of your computing system. These emails are often from an unknown sender or from known senders whose security has been compromised.

All real-world systems have errors or design flaws that can be exploited to compromise them. Regular software updates help fix errors that could compromise a computing system. Along with updates, users should review the permission settings of programs to protect their privacy and understand what user information the program may or may not be collecting.

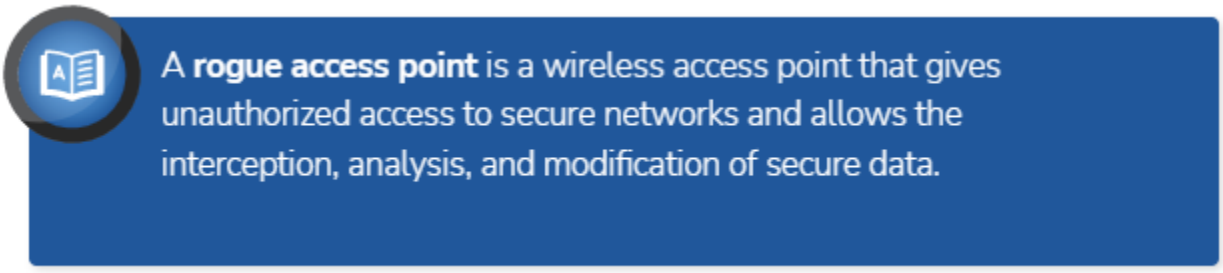
Phishing, Keylogging and Rogue Access Points



That personal information can then be used to access sensitive online resources, such as bank accounts and emails. Phishing normally takes place in the form of a phony email or website and preys on the user's fear. Have you ever received an email from "your bank" saying that your account may be in danger and that you should "click here" to reset your password? That is a phishing scam.



Keyloggers are often found on public computers or added to a user's computer unknown to them. Once the keylogger is active, the installer of the program can see every website you have visited and every password you have typed. They can access your bank account, social media accounts, school email, and everything else for which you use a password. In general, a good rule of thumb is to stay away from a public computer. And chances are, you are not going to go up to a random computer in the middle of a train station and start logging in to all your accounts. But what you might be more inclined to do is use a public network. After all, your mobile data comes at a premium and you want to connect to wifi whenever possible. So although you might think, "Hmmm… should I really connect to the network: Wi Believe I Can Fi" and you decide, "Eh, what is the worst that could happen?" If you join that network, now you are at the risk of someone else accessing your personal data. Data sent over public networks can be intercepted, analyzed, and modified.



Let's say a company is having a conference meeting and is having trouble connecting to the network with their laptops. With quick thinking by a brand new IT guy, he plugs in a router into a nearby Ethernet port and tells everyone in the conference room to connect to the router. This access point is now "rogue" and is a serious security threat. Although added by a good-willed employee, this access point (AP) can be used to bypass firewalls and damage company data.

The Art of Persuasion

Information security is not just focused on *technical* attacks on computer systems. Many malicious attackers use *social engineering* techniques as well. **Social engineering** in a computational sense refers to the psychological manipulation of people into performing actions or divulging confidential information. Below are examples of social engineering in a non-computational manner, showing how persuasive people can be.

Watch this video to learn more: https://www.youtube.com/watch?v=cFdCzN7RYbw

Complete Reflection #3 at the end of the packet

Complete the Investigate Task at the end of the packet

APCSP Summer Assignment Worksheets and Answer Sheet

All work must be shown on this answer sheet in your own handwriting unless otherwise stated. You only need to submit items on Schoology from this point forward. Schoology will be open for submission the day before the first day of class. Be sure to follow the directions on Schoology for submissions.

Summer Assignment Rubric

	Points	**Points Earned**
Part I — Encryption	**50**	
Worksheet #1	10	
Worksheet #2	10	
Make Your Own Cipher	30	
Part II — Cybersecurity	**50**	
Reflection #1	2	
CIA Components Multiple Choice Questions	3	
Reflection #2	2	
Reflection #3	2	

Investigate	41	
Total	**100**	

Part 1–Encryption

Worksheet #1: You must show **ALL** your work for any **credit**. Work must be in your own ***handwriting*** to receive **credit**. Each problem is one point.

Encode the following messages:

1. Caesar cipher with shift +3

   HELLO WORLD

2. Caesar cipher with shift +12

   WAFFLE FRIES

Decode the following messages:

3. Caesar cipher with shift +5

   TSAFH FYNTS

4. Caesar cipher with shift +21

   YMNOM VIBZ

5. Caesar cipher using frequency analysis. Shift is _____.

   KBKXE UTK

6. Caesar cipher using frequency analysis. Shift is _____.

   KZBSV CRYGO BCLBS XQWKI PVYGO BC

7. Caesar cipher using frequency analysis. Shift is _____.

KGYEZ UHXKG Q

8. Caesar cipher using frequency analysis. Shift is _____.

EWTNK PIKUO AHCXQ TKVGU RQTV

9. Caesar cipher using frequency analysis. Shift is _____.

EXZHX QELKO BXAV

10. Caesar cipher with shift +14 = -12

QWDVS FFCMO ZHM

Worksheet #2: Practice using Matrix Transposition ciphers by encoding and decoding the messages, You must show **ALL** your work for any **credit**. Work must be in your own ***handwriting*** to receive **credit**. Each problem is worth two points.

Encode the following messages. Be sure to use the proper format:

1. Matrix Transposition with rows of 4:

   THE PACKAGE WILL ARRIVE AT EIGHT

2. Matrix Transposition with rows of 5:

   WE WILL HAVE PIZZA AND CANDY

3. Matrix Transposition using the keyword MOUSE:

    WE LOVE COMPUTER SCIENCE PRINCIPLES

Decode the following messages:

4. Basic Matrix Transposition:

    IBHPL TWEEU AWIIC TBOLN OEAZL TMRTZ

5. Matrix Transposition with keyword GRAPE:

    EFAAO EOZSU RRYTB YEALD

Worksheet #3: Create Your Own Cipher

The two ciphers discussed so far both encode messages effectively, but are rather simple to crack and therefore are not very secure. Your final task is to come up with your own algorithmic method for encoding a message. Your method might be based on Caesar or Matrix Transposition, something completely different, or even a combination of methods. Complete the following:

a. Write a brief description of your encryption algorithm – one to two paragraphs, enough to understand how it works. (10 points)

b. Encode the plaintext message below using your method and write out the ciphertext. (5 points)

THIS IS A HIDDEN MESSAGE

c. Show how to decode the message back to the plaintext. (5 points)

d. Create a short message of your own and encode it using your method. Give the just ciphertext of this method so that it can be decoded when we come back to school! (10 points)

Obviously computers are able to utilize significantly more complex encryption methods, encode and decode *much* larger amounts of data than we could ever do by hand, and do it all in fractions of a second. I hope that this exercise, and the associated readings, have given you an idea of what is possible using today's computing capabilities. We will discuss more about data encryption, with its benefits and drawbacks, in the first unit of the course. See you in September!!

Part 2– Cybersecurity

Reflection #1: What is the difference between symmetric and asymmetric communication? (2 Points)
Submit your answer on Schoology *after September 1st*.

In the following scenarios, determine which of the **CIA components** is breached (1 point each):

1. Alice is buying books from an online retail site. She finds that she is able to change the price of a book from 19.99 to 1.99.
    a. Confidentiality
    b. Integrity
    c. Availability
2. Kim takes her college admissions test and is waiting to get her results by email. By accident, Kim's results are sent to Karen.
    a. Confidentiality
    b. Integrity
    c. Availability
3. Rob opens his fitness tracking app to start logging a workout. The app crashes, and he is unable to log his workout.
    a. Confidentiality
    b. Integrity
    c. Availability

Reflection #2: After watching the authenticity video, explain areas in your life where you may have to use multifactor authentication. Give specific examples. (2 points)

Submit your answer on Schoology *after September 1st*.

Reflection #3: Explain how computing resources can be protected and how computing resources can be misused. Provide specific examples with your explanations. (2 Points)

Submit your answer on Schoology *after September 1st*.

InvestigateTask (41 points total)

Create a presentation (PowerPoint or Google Slides) on all of these examples of malware/social engineering. There should be one slide for each of the 11 examples.

- baiting
- dumpster diving
- shoulder surfing
- tailgating
- pretexting
- diversion theft
- backdoors
- Trojan horses
- time bombs
- spyware
- botnets

Include the following information in your slide for each of the 11 examples:

1. What is the malware/social engineering example?
2. How does it work?
3. What are some real-world examples of it?
4. How does this impact confidentiality, integrity, availability, or authenticity?
5. How is it prevented and/or removed?

6. Have you experienced this type of social engineering? How and what happened?

Submit your presentation as a *PDF* on Schoology.

***Only PDFs* will be accepted and graded.**