

Project 4: Build It

Design Approach

The Secure Gradebook I built is divided into three programs: setup.c, gradebookadd.c, and gradebookdisplay.c. The Gradebook is stored as a C struct, with lists of Students and Assignments (which are also C structs). All of the structs are defined in gradebookobjects.h.

- Assignment struct

```
typedef struct assignment {  
    char name[50];  
    int points;  
    double weight;  
    int grade;  
} Assignment;
```

- Student struct

```
typedef struct student {  
    char first_name[50];  
    char last_name[50];  
    int num_assignments;  
    Assignment assignments[100];  
} Student;
```

- Gradebook struct

```
typedef struct unpadded_gradebook {  
    char name[50];  
    int num_assignments;  
    Assignment assignments[100];  
    int num_students;  
    Student students[500];  
} unpadded_gradebook;  
  
typedef struct gradebook {  
    char name[50];  
    int num_assignments;  
    Assignment assignments[100];  
    int num_students;  
    Student students[500];  
    char padding[64 - (sizeof(unpadded_gradebook) % 64)];  
} Gradebook;
```

Note: The gradebook struct includes padding so that it can be properly encrypted (this will be described later in this write-up)

Each of the structs above use fixed lengths for any array. *The following assumptions were made:*

- Assignment names are limited to a maximum of 50 characters
- Student first names are limited to a maximum of 50 characters

- Student last names are limited to a maximum of 50 characters
- Each student can have a maximum of 100 assignments
- Gradebook filenames are limited to a maximum of 50 characters
- Each gradebook can have a maximum of 100 assignments
- Each gradebook can have a maximum of 500 students

These array lengths are fixed because it allows the program to validate any input and prevent buffer overflows, which will be described later in this write-up).

The TAs and Professors can create a new gradebook using `setup.c`, modify an existing gradebook using `gradebookadd.c`, and read an existing gradebook using `gradebookdisplay.c`. Here is how each of the programs work internally:

1. `setup.c`

- If a valid filename is provided, and a gradebook with the provided filename does not already exist in the current directory, **the program opens a new file with the provided filename.**
- The program creates a new, empty instance of the Gradebook struct: `new_gb`.** The new instance of the Gradebook contains 0 assignments and 0 students.
- The program generates a random 128-bit IV** using openssl's `RAND_bytes` function.
- The program generates a random 128-bit encryption key** using openssl's `RAND_bytes` function.
- The program generates a random 128-bit MAC key** using openssl's `RAND_bytes` function.
- The program prints a 256-bit random key to the user** (the first 128 bits are the encryption key, and the second 128 bits are the MAC key).
- The program encrypts `new_gb` (the Gradebook struct)** with the encryption key and IV and stores it in a buffer called `ciphertext`. Encryption is done using openssl functions.
- The program runs the MAC algorithm with the randomly generated MAC key and produces a tag for the Gradebook encryption** using openssl's MAC function and `EVP_md5` hash. The tag is 128-bits in length.
- The program writes the ciphertext (the encrypted Gradebook) to the new gradebook file.**
- The program writes the IV to the new gradebook file.**
- The program writes the tag produced by MAC to the gradebook file.**
- The program writes the length of the ciphertext to the gradebook file.**

2. `gradebookadd.c`

- If a valid filename is provided, and a gradebook with the provided filename exists in the current directory, **the program opens the file with the provided filename.**
- The program reads the following information from the gradebook file:**
 - Ciphertext
 - IV

- iii. MAC tag
- iv. Ciphertext length
- c. **The program reads the provided 256-bit key from the command line**, and splits it into two parts (128-bit encryption key and 128-bit MAC key).
- d. **The program runs the MAC algorithm with the MAC key from the command-line and produces a tag for the ciphertext** using openssl's MAC function and EVP_md5 hash. The tag is 128-bits in length.
- e. **The program compares the MAC tag from the gradebook file to the MAC tag that was generated with the ciphertext.** If they are the same, the gradebook file was NOT modified.
- f. **The program decrypts the ciphertext** with the ciphertext, ciphertext length, encryption key, and IV using openssl functions. **The program stores the decryption into a Gradebooks struct.**
- g. **The program reads the rest of the inputs from the command line** and determines if a valid combination has been provided.
- h. **The program performs one of the following actions by modifying the gradebook:**
 - i. Add assignment
 - 1. The program creates a new instance of the Assignment struct with the user's provided inputs and adds it to the Gradebook assignments array. The program increments the Gradebook num_assignments property by 1.
 - 2. The program adds new instances of the Assignment struct with the user's provided inputs to each of the students' assignments array. Each of the students' num_assignments property is increased by 1. Each of the students' grades on the assignment is initialized to 0.
 - ii. Delete assignment
 - 1. The program goes through the Gradebook assignments array and looks for an assignment with the same name that the user provided. If an assignment with the same name exists, it is removed from the Gradebook by shifting all following assignments up by one index and decrementing the Gradebook num_assignments property.
 - 2. The program goes through each students' assignments array and looks for an assignment with the same name that the user provided. If an assignment with the same name exists, it is removed from the students' assignments array by shifting all following assignments up by one index and decrementing the students' num_assignments property.
 - iii. Add student
 - 1. The program creates a new instance of the Student struct with the user's input and adds it to the Gradebook students array. The Student struct is initialized with all of the assignments that are in

the Gradebook assignments array (with a starting grade of 0). The program increments the Gradebook num_students property.

iv. Delete student

1. The program goes through the Gradebook students array and looks for a student with the same first and last name that the user provided. If a student with the same name exists, it is removed from the Gradebook's students array by shifting all following students up by one index and decrementing the Gradebook's num_students property.

v. Add grade

1. The program goes through the Gradebooks students array and looks for a student with the same first and last name that the user provided. If a student with the same name exists, the program goes through the student's assignment array and looks for an assignment with the same name that the user provided. If an assignment with the same name exists, the student's grade on the assignment is updated with the user's input.

- i. After performing the action and modifying the gradebook, **the program opens a fresh file with the same name as the original gradebook.**
 - j. **The program generates a random 128-bit IV** using openssl's RAND_bytes function.
 - k. **The program encrypts the Gradebook struct** with the encryption key from the command-line and randomly-generated IV and stores it in a buffer called ciphertext. Encryption is done using openssl functions.
 - l. **The program runs the MAC algorithm with the MAC key from the command-line and produces a tag for the Gradebook encryption** using openssl's MAC function and EVP_md5 hash. The tag is 128-bits in length.
 - m. **The program writes the ciphertext (the encrypted Gradebook) to the gradebook file.**
 - n. **The program writes the randomly-generated IV to the gradebook file.**
 - o. **The program writes the tag produced by MAC to the gradebook file.**
 - p. **The program writes the length of the ciphertext to the gradebook file.**
3. gradebookdisplay.c
- a. If a valid filename is provided, and a gradebook with the provided filename exists in the current directory, **the program opens the file with the provided filename.**
 - b. **The program reads the following information from the gradebook file:**
 - i. Ciphertext
 - ii. IV
 - iii. MAC tag
 - iv. Ciphertext length
 - c. **The program reads the provided 256-bit key from the command line**, and splits it into two parts (128-bit encryption key and 128-bit MAC key).

- d. **The program runs the MAC algorithm with the MAC key from the command-line and produces a tag for the ciphertext** using openssl's MAC function and EVP_md5 hash. The tag is 128-bits in length.
- e. **The program compares the MAC tag from the gradebook file to the MAC tag that was generated with the ciphertext.** If they are the same, the gradebook file was NOT modified.
- f. **The program decrypts the ciphertext** with the ciphertext, ciphertext length, encryption key, and IV using openssl functions. **The program stores the decryption into a Gradebook struct.**
- g. **The program reads the rest of the inputs from the command line** and determines if a valid combination has been provided.
- h. **The program performs one of the following actions by reading the gradebook:**
 - i. Print assignment
 - 1. The program goes through the Gradebook assignments array to check that an assignment with the provided name exists in the Gradebook.
 - 2. The program copies all of the Gradebook students to a different array.
 - 3. If the user provided the -A flag
 - a. The program sorts the copied students array by last name, first name
 - 4. If the user provided the -G flag
 - a. The program sorts the copied students array by grade
 - 5. For each student in the copied students array, the program goes through the student's assignments array and locates the assignment with the name the user provided. The program prints the students last name, first name, and grade on the assignment to a new line.
 - ii. Print student
 - 1. The program goes through the Gradebook's students array to find a student with the user provided first and last names. If a student with the provided name exists, the program goes through each assignment in the student's gradebook and prints the assignment name and grade on a new line.
 - iii. Print final
 - 1. The program copies all of the Gradebook students to a different array.
 - 2. If the user provided the -A flag
 - a. The program sorts the copied students array by last name, first name
 - 3. If the user provided the -G flag
 - a. The program sorts the copied students array by final grade

4. The program goes through the copied students array and prints each student's last name, first name, and final grade on a new line.
 - i. After performing the action, **the program closes the gradebook.**

While implementing the specifications of this project, *the following assumptions were made*:

- If a flag expects an input such as (-AN <assignment name>), and no such input is provided, the program will exit and print "invalid".
- If extra options are provided, but the required options are provided, then the extra options will be ignored.

Attacks

This secure gradebook program protects against the following attacks.

Attack 1: Read From Cloud

In this attack, a student would like to see the grades of all the other students in the class. The student is able to gain access to the cloud where the course gradebook is stored. They obtain the gradebook file, and the student is able to read the contents of the file. However, reading the contents of the file does **NOT** tell the student any useful information. When the student opens the file, they will see the encryption of the gradebook:

```

...O...[...Pr.p.b.'X<...b
j.....vM..4.....<.>..@Z
...,W..S..Bd.'...S3PX..
.....t~S#...x.L+S.....^..
....j.....S.N...3SB.!=..
.....&."-..Bo.....
.._..yHF.>.....c..{..
pe.....|.....;.....
bs.....,1.....8}.E...
.|....."9.Z.....M*....
..BPh..l.T.....g.2.....
....u....2.e...v....vb.
yS.5.....9,.G9^.=F*L.Ow<
..E..<..>25 ]aYK..E_...
...r.....F.jU...|. &+.
..B.5.....LR.Yy.#
.....r:....._(+)...
Z..d0.n^7.N1)....RqAqQ+
.r...hN.....C...*...&.LJ
o....'s.V:....U.....ag
|..."L%....AT3.Q.t.[.c.
t^1.....H.....^bW.ZX".
.A/..C.._...iA.xYp..zw.
Am..5v..(A_!....)].."...

```

This program guarantees the privacy of the gradebook data by using AES encryption.

Before being written to the gradebook file, the gradebook is encrypted with a 128-bit key (which only the TAs and Professors have access to) and a random 128-bit IV (which is included in the gradebook file). The 128-bit key is securely and randomly generated in setup.c using rand_bytes and printed for the TA and Professor to see. The 128-bit IV is securely and randomly generated each time the gradebook is modified. The lines of code that encrypt the gradebook are:

- In setup.c: lines 247-248
- In gradebookadd.c: lines 580-581

If the student would like to view the contents of the gradebook using the gradebookdisplay.c program, they must provide the correct key. If the student provides an incorrect key, as seen below, the program will exit and return “invalid”.

```

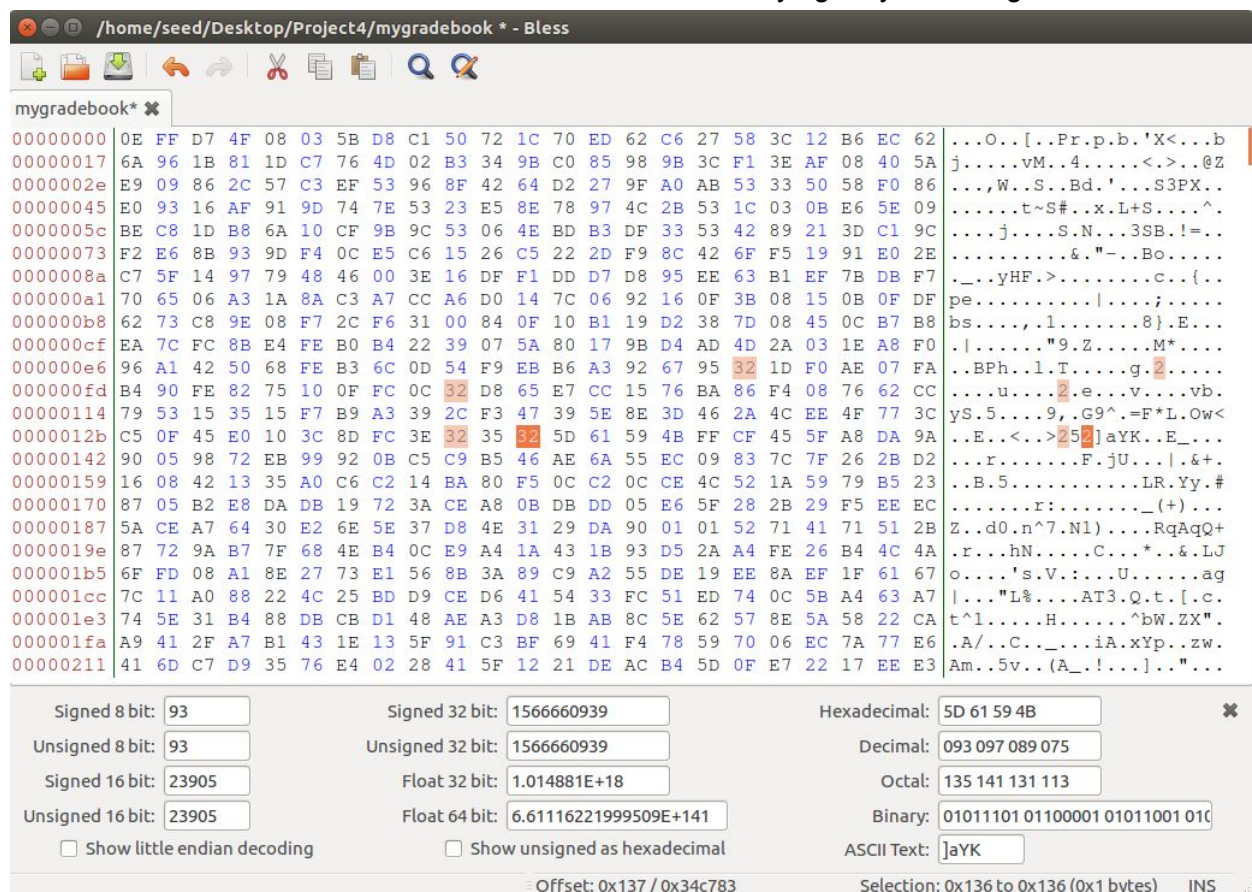
[11/27/20]seed@VM:~/.../Project4$ gradebookdisplay -N mygradebook -K BFE9F8FA2E03C438F3C973D1C05394E46F5004479228
909F71A99BF82F1C34D8 -PA -AN Midterm2 -A
invalid

```

AES-128 encryption is considered secure, so it will take a significant amount of time for any attacker to brute force the key. The amount of time to brute force the correct key is so long that an attacker would not be able to perform such an attack.

Attack 2: Modify the Gradebook

In this attack, a student who got a bad grade on the midterm would like to change their grade. The student is able to gain access to the cloud where the course gradebook is stored. They obtain the gradebook file, and the student is able to change the contents of the file. However, this attack will not be successful. If the user modifies even one byte of the gradebook file, when the professor or TA runs the gradebookadd or gradebookdisplay programs, the program will exit and return “invalid”. Here is a screenshot of a student modifying a byte in the gradebook file:



Here is a screenshot of what happens when the professor or TA try to read the modified gradebook:

```
[11/27/20]seed@VM:~/.../Project4$ gradebookdisplay -N mygradebook -K BFE9F8FA2E03C438F3C973D1C05394E46F5004479228
909F71A99BF82F1C34D7 -PA -AN Midterm2 -A
invalid
```


This program guarantees the integrity of the gradebook data by using the openssl HMAC function. Before the encryption of the gradebook is written to the gradebook file, the program generates a tag of the encryption with a 128-bit key (which only the TAs and Professors have access to) and the EVP_md5() hash. The encryption and tag are then written to the gradebook file. The 128-bit key is securely and randomly generated in setup.c using rand_bytes and printed for the TA and Professor to see. The lines of code that generate the tag are:

- In setup.c: line 253
- In gradebookadd.c: line 586

Whenever the gradebookadd.c and gradebookdisplay.c programs read from a gradebook file, the programs use the openssl HMAC function to generate a tag with the ciphertext from the gradebook file and the key that the TA or professor provides. If the tag that gradebookadd.c and gradebookdisplay.c generates is the same as the tag in the gradebook file, then the gradebook file was not tampered with. However, if the tags are different, then the gradebook file was altered and cannot be trusted, so the program exits and returns “invalid”. The lines of code that compare the tags are:

- In gradebookadd.c: lines 276-280
- In gradebookdisplay.c: lines 271-275

Attack 3: Buffer Overflow

In this attack, the attacker tries to alter the behavior of the programs by overwriting pointers on the stack. In this case, the attacker was able to identify where on the stack the return register is located using gdb. The attacker would like to overwrite the return register so it launches some other function (which is not the expected behavior of this program). The attacker would like to overwrite the return register using the input that each of the programs take. The following options take in an input that is written to a buffer in the program:

- -N <filename>
- -K <key>
- -AN <assignment name>
- -FN <first name>
- -LN <last name>

Each of the buffers which store these inputs are different lengths. The lengths are as follows:

- -N <filename>
`char name[50];`
- -K <key>
`unsigned char decrypt_key[32];`
- -AN <assignment name>
`char name[50];`
- -FN <first name>
`char first_name[50];`

- -LN <last name>

```
char last_name[50];
```

When reading the user's inputs from the command line, the program checks that the lengths of the input are valid before writing the input into the program buffers. If the lengths of the inputs are too long, the program exits and returns "invalid".

If the attacker attempts to overwrite the return register by inputting data that is longer than the buffer, the program will exit and the attack will be unsuccessful. This can be seen in the following cases:

```
[11/27/20]seed@VM:~/.../Project4$ setup -N ThisInputIsWayTooLongToBeAcceptedByTheGradebookPrograms
invalid

[11/27/20]seed@VM:~/.../Project4$ gradebookdisplay -N mygradebook -K BFE9F8FA2E03C438F3C973D1C05394E46F5004479228
909F71A99BF82F1C34D71234 -AA -AN Project4 -P 100 -W 0.01
invalid

[11/27/20]seed@VM:~/.../Project4$ gradebookdisplay -N mygradebook -K BFE9F8FA2E03C438F3C973D1C05394E46F5004479228
909F71A99BF82F1C34D7 -AA -AN ThisInputIsWayTooLongToBeAcceptedByTheGradebookPrograms -P 100 -W 0.01
invalid

[11/27/20]seed@VM:~/.../Project4$ gradebookdisplay -N mygradebook -K BFE9F8FA2E03C438F3C973D1C05394E46F5004479228
909F71A99BF82F1C34D7 -AS -FN ThisInputIsWayTooLongToBeAcceptedByTheGradebookPrograms -LN Ford
invalid

[11/27/20]seed@VM:~/.../Project4$ gradebookdisplay -N mygradebook -K BFE9F8FA2E03C438F3C973D1C05394E46F5004479228
909F71A99BF82F1C34D7 -AS -FN Maddie -LN ThisInputIsWayTooLongToBeAcceptedByTheGradebookPrograms
invalid
```

The code for checking the lengths of all the inputs is:

- -N <filename>: in setup.c (line 178), in gradebookadd.c (line 231)
- -K <key>: in gradebookadd.c (lines 239-242), in gradebookdisplay.c (lines 234-237)
- -AN <assignment name>: in gradebookadd.c (lines 37-46), in gradebookdisplay.c (lines 61-70)
- -FN <first name>: in gradebookadd.c (lines 49-58), in gradebookdisplay.c (lines 73-81)
- -LN <last name>: in gradebookadd.c (lines 49-58), in gradebookdisplay.c (lines 73-81)

Attack 4: Invalid Input

In this attack, the attacker attempts to input invalid data to get an unexpected output. Each of the options require different data types as specified below:

- -N
 - may be specified with a string of alphanumeric characters (including underscores and periods)
- -AN
 - alphabetic characters (a-z, A-Z) in upper and lower case and numbers (0-9)
 - may not contain spaces
 - case sensitive
- -FN
 - alphabetic characters (a-z, A-Z) in upper and lower case
 - may not contain spaces

- case sensitive
- -LN
 - alphabetic characters (a-z, A-Z) in upper and lower case
 - may not contain spaces
 - case sensitive
- -P
 - required to be a non-negative integer
- -W
 - required to be a real number in [0,1]
- -G
 - required to be a non-negative integer

In this attack, the attacker wants some unexpected output, so they will input data that does not meet the requirements above. An attack like this could result in a segmentation fault or some other corruption of the gradebook.

However, this attack will be unsuccessful on this gradebook system, because I implemented whitelisting. Whitelisting defines a set of valid characters. Each of the options above have a corresponding function that guarantees the input only contains valid characters. The lines of these functions are:

- -N: in setup.c (lines 23-32), in gradebookadd.c (lines 25-34), in gradebookdisplay.c (49-58)
- -AN: in gradebookadd.c (lines 37-46), in gradebookdisplay.c (lines 61-70)
- -FN: in gradebookadd.c (lines 49-58), in gradebookdisplay.c (lines 73-82)
- -LN: in gradebookadd.c (lines 49-58), in gradebookdisplay.c (lines 73-82)
- -P: in gradebookadd.c (lines 61-67)
- -W: in gradebookadd.c (lines 70-82)
- -G: in gradebookadd.c (lines 61-67)

Each of the above functions are called before performing any action or storing the input to any variables in the program. This can be seen in lines 365-565 in gradebookadd.c and lines 336-486 in gradebookdisplay.c. If the program reads any input that does not meet the requirements above, the program exits and returns “invalid”.

An unsuccessful attempt of this attack can be seen in the following screenshots:

```
[11/27/20]seed@VM:~/.../Project4$ setup -N grades@1
invalid
```

```
[11/27/20]seed@VM:~/.../Project4$ gradebookadd -N mygradebook -K 51185E1BA712573683646C9B48E758C9C5559E6126033F0E
8E38884E2E1EBA7E -AA -AN Final! -P 100 -W 0.2
invalid
```

```
[11/27/20]seed@VM:~/.../Project4$ gradebookadd -N mygradebook -K 51185E1BA712573683646C9B48E758C9C5559E6126033F0E
8E38884E2E1EBA7E -AS -FN Maddie! -LN Ford
invalid
```

```
[11/27/20]seed@VM:~/.../Project4$ gradebookadd -N mygradebook -K 51185E1BA712573683646C9B48E758C9C5559E6126033F0E
8E38884E2E1EBA7E -AS -FN Maddie -LN Ford!
invalid
```

```
[11/27/20]seed@VM:~/.../Project4$ gradebookadd -N mygradebook -K 51185E1BA712573683646C9B48E758C9C5559E6126033F0E8E38884E2E1EBA7E -AA -AN Final -P a -W 0.2  
invalid
```

```
[11/27/20]seed@VM:~/.../Project4$ gradebookadd -N mygradebook -K 51185E1BA712573683646C9B48E758C9C5559E6126033F0E8E38884E2E1EBA7E -AA -AN Final -P 100 -W 0.0.2  
invalid
```

```
[11/27/20]seed@VM:~/.../Project4$ gradebookadd -N mygradebook -K 51185E1BA712573683646C9B48E758C9C5559E6126033F0E8E38884E2E1EBA7E -AG -FN Ted -LN Mason -AN Final -G b  
invalid
```