

Entwerfen und Implementieren einer Verschlüsselungssoftware

Ausarbeitung

Bachelor of Science

Studiengang Informationstechnik

an der

Dualen Hochschule Baden-Württemberg Karlsruhe

von

Nico Schrodt

Abgabedatum 27. April 2022

Bearbeitungszeitraum	5 + 6 Semester
Kurs	TINF19B3
Dozent	Daniel Lindner

Inhaltsverzeichnis

1	Einführung	1
1.1	Ziel der Arbeit	1
1.2	Repository	1
1.3	Commit-Historie	1
1.4	UML-Diagramm	2
2	Clean Architecture	3
2.1	Geplante Schichtenarchitektur	3
2.2	Umsetzung	3
2.2.1	Benutzeroberfläche	3
2.2.2	Verschlüsselungsdienst	3
3	Entwurfsmuster	4
4	Programming Principles	5
4.1	SOLID	5
4.1.1	Single Responsibility Principle	5
4.1.2	Open/Closed Principle	5
4.1.3	Liskov Substitution Principle	6
4.1.4	Interface Segregation Principle	6
4.1.5	Dependency Inversion Principle	6
4.2	GRASP	7
4.2.1	Low Coupling	7
4.2.2	High Cohesion	7
4.3	DRY	7
5	Refactoring	8
5.1	Code Smells	8
5.1.1	Code Smells 1 Duplicated Code	8
5.1.2	Code Smells 2 Long Method	9
5.1.3	Code Smells 3 Large Class	9
5.2	Angewendete Refactorings	10
5.2.1	Refactoring 1 Extract Method	10
5.2.2	Refactoring 2 Replace Error Code with Exception	10
6	Unit Tests	12
6.1	Verwendete Unit Tests und getesteter Code	12
6.1.1	Mocks	12
6.1.2	Code Coverage	13

6.2	Anwendung der ATRIP-Regeln	14
-----	--------------------------------------	----

1 Einführung

Dieses Kapitel befasst sich vorwiegend mit relevanten Grundlagen der Arbeit. Unter anderem wird das Ziel spezifiziert, elementare Aspekte der Arbeitsweise eines Prozessors werden erläutert und die verschiedenen Werkzeuge mit denen das Ziel realisiert wird werden aufgeführt.

1.1 Ziel der Arbeit

In dieser Arbeit soll ein Simulationsprogramm geschrieben werden, mit dem mehrere unterschiedliche 8-Bit Prozessoren simuliert werden können. Dazu sollen die Grundlegenden Eigenschaften in kurzen Lernprogrammen erläutert werden. Ebenfalls soll es eine interaktive Einweisung geben wie der Simulator verwendet werden kann.

1.2 Repository

Der Quellcode kann in folgendem GitHub-Repository abgerufen werden:

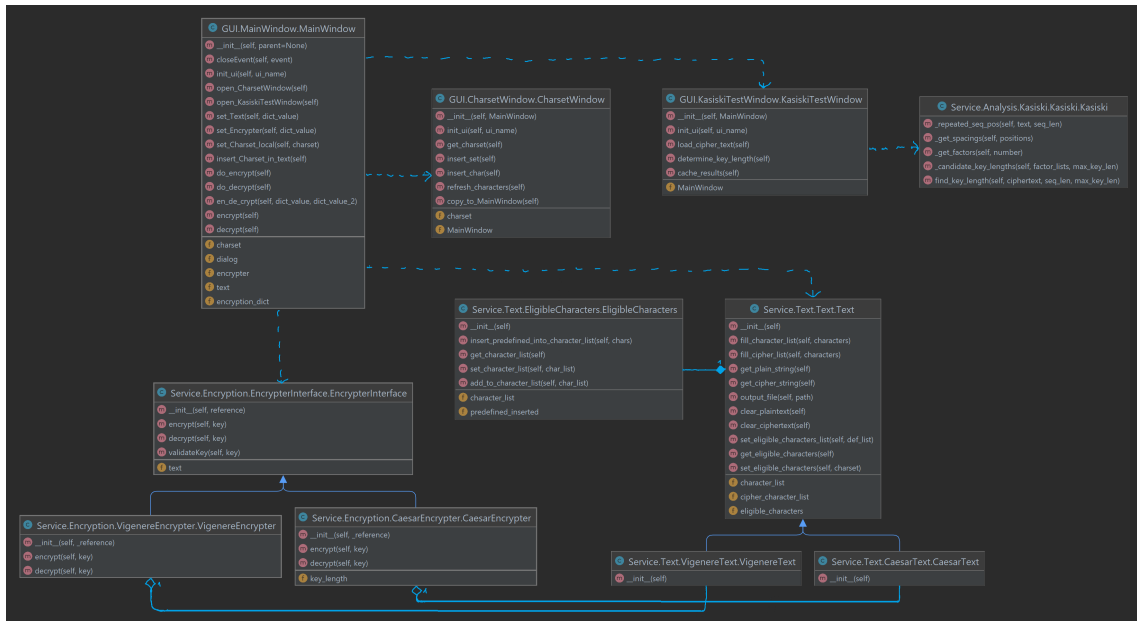
<https://github.com/NicoSchrodt/EncryptionService>

1.3 Commit-Historie

Der Titel jedes Kapitels hat eine Fußnote die den Commit angibt während dem das Kapitel geschrieben wurde bzw. dieses zuletzt verändert wurde. Für Abschnitte in denen auf spezifische Änderungen eingegangen wird, werden i.d.R eigene Commits angegeben mit direktem Link (100% RickRoll-Frei, für paranoide ist der Commit auch separat angegeben).

Sollte dieser Satz noch hier sein: Dies ist eine Vorabversion als Notfalls falls ich vergessen habe die fertige Version rechtzeitig hochzuladen.

1.4 UML-Diagramm



2 Clean Architecture

Der Sinn einer Clean Architecture ist es das Programm in klar definierte Schichten zu zerlegen die unabhängig voneinander ausgetauscht werden können. Dadurch soll idealerweise die Langlebigkeit und Wartbarkeit eines Projekts gewährleistet werden können.

2.1 Geplante Schichtenarchitektur

Für dieses Projekt sind zwei Schichten vorgesehen. Einmal die Benutzeroberfläche (GUI) welche mit Qt implementiert wird und die Logik, welche unter anderem die Verschlüsselung vornimmt. Der Benutzer soll ausschließlich mit der von Qt generierten Oberfläche interagieren z.B. durch Textfelder oder Knöpfe, welche vorkonfigurierte Befehle ausführen.

2.2 Umsetzung

Platzhalter

2.2.1 Benutzeroberfläche

Platzhalter

2.2.2 Verschlüsselungsdienst

Platzhalter

3 Entwurfsmuster

Das für den Umfang dieses Programmentwurfs verwendete Entwurfsmuster ist der Dekorierer. Aufgabe des Dekorierers ist es eine Klasse oder Funktion um einen oder mehrere Aspekte zu erweitern ohne die Klasse selbst zu verändern. Das Entwurfsmuster wurde in der Klasse 'EncrypterInterface.py' angewendet.

```
class EncrypterInterface:
    def __init__(self, reference):
        self.text = reference

    def encrypt(self, key):
        # Encrypt the character list in text-object
        pass

    def decrypt(self, key):
        # Decrypt the character list in text-object
        pass
```

'EncrypterInterface' dient wie der Name bereits verrät als Interface für konkrete Encrypter. Dabei ist aber nicht gewährleistet, dass die konkrete Implementierung die im Interface beschriebenen Funktion selbst implementiert. Der Dekorierer soll hier die Aufgabe übernehmen, auf eine konkrete Implementierung zu kontrollieren und bei Fehlen dieser eine Exception auszulösen.

```
class EncrypterInterface(metaclass=abc.ABCMeta):
    def __init__(self, reference):
        self.text = reference

    @abc.abstractmethod
    def encrypt(self, key):
        # Encrypt the character list in text-object
        raise NotImplementedError

    @abc.abstractmethod
    def decrypt(self, key):
        # Decrypt the character list in text-object
        raise NotImplementedError
```

Die Funktion des Dekorierers beschränkt sich hier auf konkrete Implementierungen des EncrypterInterfaces, Also Klassen die von 'EncrypterInterface' erben. Das Interface selbst könnte potentiell immer noch instantiiert und verwendet werden ohne die Exception auszulösen.

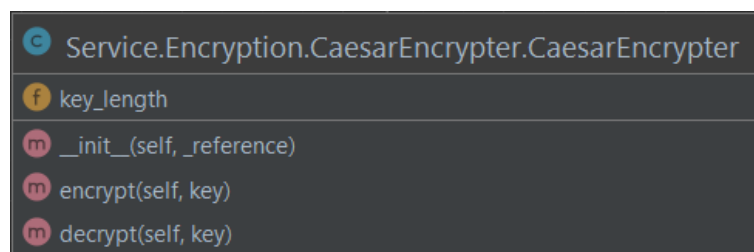
4 Programming Principles

In diesem Abschnitt werden kurz einige Programming Principles erläutert und deren Anwendung an Beispielen in diesem Projekt aufgezeigt.

4.1 SOLID

4.1.1 Single Responsibility Principle

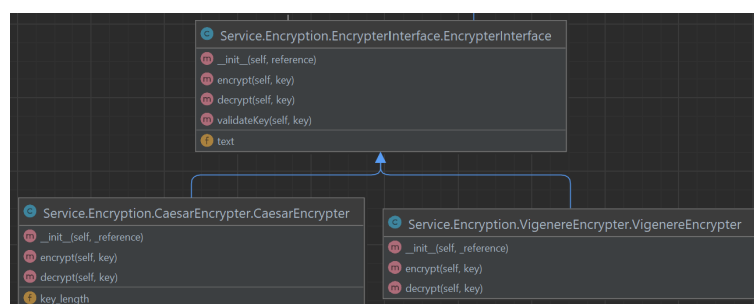
Das Single Responsibility Principle steht für die Anforderung das jede Klasse nur eine einzige Aufgabe bzw. Verantwortung haben soll. Sinn dahinter ist es Komplexität und unerwünschte Kopplung zu vermeiden. Generell ist nämlich davon auszugehen das eine Klasse mit mehreren Verantwortungen Interaktionen zwischen diesen hat, was unter anderem das Ändern einzelner erschwert. Als Beispiel dafür wird in der unteren Abbildung eine konkrete Implementierung der Encrypter Klasse hergezogen.



Die Klasse hat effektiv eine Aufgabe. Sie erhält bei Instanziierung ein Textobjekt. Auf diesem Textobjekt werden Verschlüsselungen durchgeführt, dabei wird lediglich zwischen Ver- und Entschlüsseln unterscheiden.

4.1.2 Open/Closed Principle

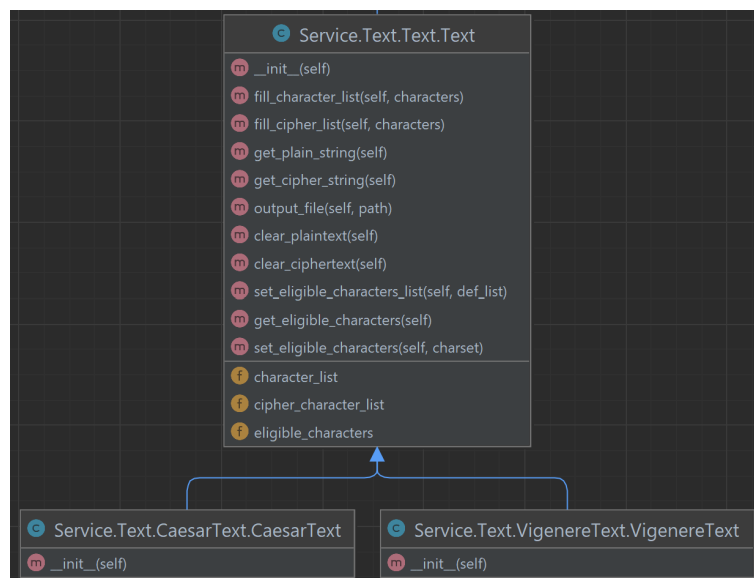
Das Open/Closed Principle beschreibt das Designziel Klassen, Funktionen, etc. so aufzubauen das sie offen sind für Erweiterungen und geschlossen für Veränderungen. Konkret heißt das, neue Anforderungen sollen eher durch z.B. Vererbung realisiert werden, statt konkreten Modifikationen in der relevanten Klasse. Dieses Prinzip wurde für das Projekt beispielsweise durch Vererbung realisiert. Statt einer generellen Encrypter-Klasse wird ein Interface verwendet, welches für jede verschiedene Anwendung eine davon erbende Klasse hat.



Dadurch ist es möglich weitere Verschlüsselungen hinzuzufügen ohne das Interface oder die erbbenden Klassen zu modifizieren.

4.1.3 Liskov Substitution Principle

Das Liskov Substitution Principle vermittelt das Prinzip, das jede Spezialisierung, z.B. durch Polymorphie bei Vererbung, an jeder Stelle verwendet werden können muss an der auch die Generalisierung verwendet wird. Beispielsweise soll also die Funktion einer Erbbenden Klasse nicht zu einem Fehler führen an einer Stelle an der die Funktion der Ursprungs Klasse funktioniert hat.



Die Struktur an der dieses Prinzip Anwendung findet wäre in diesem Projekt die Text-Klasse sowie den Kindern dieser. Generell haben diese die selben Funktionen und können auch anstelle der Textklasse verwendet werden. Die Unterschiede liegen lediglich in der Instanziierung der Unterklassen, welche bspw. das Charset verändern.

4.1.4 Interface Segregation Principle

Mit dem Interface Segregation Principle soll verhindert werden, dass Klassen ein über-spezifiziertes Interface verwenden. Ein verwendetes Interface soll also möglichst schlank sein und nicht zu viele Funktionen auf einmal anbieten. Damit soll verhindert werden, dass Klassen Zugriff auf Funktionen haben die sie gar nicht verwenden.

4.1.5 Dependency Inversion Principle

Das Dependency Inversion Principle beschreibt effektiv das Prinzip der Entkoppelung. Klassen auf einer höheren Ebene bspw. der Logik eines Programms solle nicht

von niedrigeren Klassen z.B. Benutzerinterfaces abhängen. Dies wird unter anderem durch Verschieben der Abhängigkeit erreicht, bspw. auf ein Interface und dem Übergeben von Referenzen auf konkrete Instanzen.

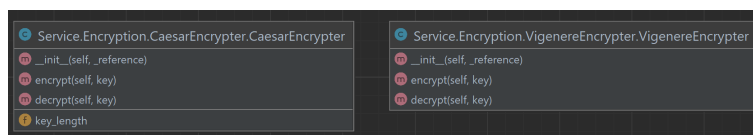
4.2 GRASP

4.2.1 Low Coupling

'Low Coupling' oder 'Niedrige Kopplung' meint das Prinzip, dass Programmcode besser ist je weniger Abhängigkeiten auf die Umgebung bestehen bzw. diese möglichst nah beieinander sind. Das soll eine flexiblere Anwendung ermöglichen und Modifikationen erleichtern, da die Anzahl an Abhängigkeiten überschaubarer ist.

4.2.2 High Cohesion

'High Cohesion' oder ein hohes Maß an Kohäsion stellt das Gegenstück zur niedrigen Kopplung dar. Hierbei soll sichergestellt werden, dass 'naher' Code tatsächlich eng miteinander arbeitet. Das kann sich z.B. darin widerspiegeln ob Code einer Klasse regelmäßig verwendet wird oder nur in Spezialfällen. Prinzipiell ist ein Weg dieses Ziel zu erreichen das Aufspalten von großen Klassen in Subklassen mit kleineren und spezifischeren Aufgaben. Im Projekt wird dies erreicht durch implementieren einzelner Klassen für jede Verschlüsselung statt einer großen 'Encrypter'-Klasse.



4.3 DRY

DRY steht für 'Don't repeat yourself'. Zentraler Angelpunkt dieses Prinzips ist das Vermeiden von Code Duplikaten, sowie das Strukturieren des Programmcodes in einer Weise das nur logisch verknüpfte Elemente sich gegenseitig beeinflussen. Oder in anderen Worten, jedes logische Konstrukt im Quellcode muss durch eine klare von anderen Aspekten getrennte Struktur repräsentiert werden. Dadurch lassen sich z.B. einige Code Smells verhindern wie "Duplicated Code" oder "SShotgun Surgery". Das Prinzip wurde angewendet hinsichtlich dem Vermeiden von Code Duplikaten, welche sich unter anderem in zahlreichen Klassen die für die Verschlüsselung zuständig sind befanden, da der Prozess der Ver- und Entschlüsselung je nach Verfahren recht ähnlich ist (bspw. Vigenère- oder Caesar-Chiffre).

5 Refactoring

Das Ziel von Refactoring ist das Verbessern der Codequalität. Für diese Arbeit ist es unterteilt in das Identifizieren von 3 verschiedenen Code Smells und das Anwenden von 2 Refactorings.

5.1 Code Smells

Unter 'Code Smells' versteht man Stellen im Programmcode, welche Verbesserungspotential aufweisen bspw. bezüglich der Übersichtlichkeit.

(**Anmerkung:** Die hier aufgelisteten Code Smells sind womöglich nicht mehr in der neuesten Version des Projekts zu finden, sondern nur noch in älteren Commits)

5.1.1 Code Smells 1 Duplicated Code

Dieses Beispiel für einen 'Duplicated Code'- Code Smells ist in der 'Caesar-Encrypter.py'-Datei zu finden. Ausschlaggebend ist hierbei, dass das Verfahren zum Ver- und Entschlüsseln effektiv gleich ist mit der Ausnahme, welcher der Starttext ist und in welche Richtung (Positiv/Negativ) der Schlüssel anzuwenden ist.

```
def encrypt(self, key):
    local_list = self.text.character_list
    local_eligible_character_list = self.text.get_eligible_characters()
    num_elg_chars = len(local_eligible_character_list)
    for i in range(len(local_list)):
        index_key = local_eligible_character_list.index(key)
        try:
            index_char = local_eligible_character_list.index(local_list[i])
            index_new = index_char + index_key
            while index_new >= num_elg_chars:
                index_new -= num_elg_chars
            self.text.cipher_character_list.append(local_eligible_character_list[index_new])
        except ValueError:
            self.text.cipher_character_list.append(local_list[i])
            if local_list[i] != " ":
                print("Invalid character '" + local_list[i] + "' found, skipped. Please add a character-set which "
                    "contains character.")
```

```
def decrypt(self, key):
    local_list = self.text.cipher_character_list
    local_eligible_character_list = self.text.get_eligible_characters()
    num_elg_chars = len(local_eligible_character_list)
    for i in range(len(local_list)):
        index_key = local_eligible_character_list.index(key)
        try:
            index_char = local_eligible_character_list.index(local_list[i])
            index_new = index_char - index_key
            while index_new < 0:
                index_new += num_elg_chars
            self.text.character_list.append(local_eligible_character_list[index_new])
        except ValueError:
            self.text.character_list.append(local_list[i])
            if local_list[i] != " ":
                print("Invalid character '" + local_list[i] + "' found, skipped. Please add a character-set which "
                    "contains character.")
```

5.1.2 Code Smells 2 Long Method

```
def __init__(self, parent=None):
    super(MainWindow, self).__init__(parent)
    self.init_ui("GUI\\MainWindow.ui")

    self.setWindowFlags(Qt.WindowType.WindowCloseButtonHint | Qt.WindowType.MSWindowsFixedSizeDialogHint)
    self.encrypter = None
    self.text = None
    self.charset = None
    self.dialog = None

    # Exit Button
    exit_button = self.button_Exit
    exit_button.clicked.connect(self.close)

    # Encrypt Button
    encrypt_button = self.button_encrypt
    encrypt_button.clicked.connect(self.encrypt)

    # Decrypt Button
    decrypt_button = self.button_decrypt
    decrypt_button.clicked.connect(self.decrypt)

    # Menubar File
    exitApplication = self.actionExit
    exitApplication.triggered.connect(self.close)

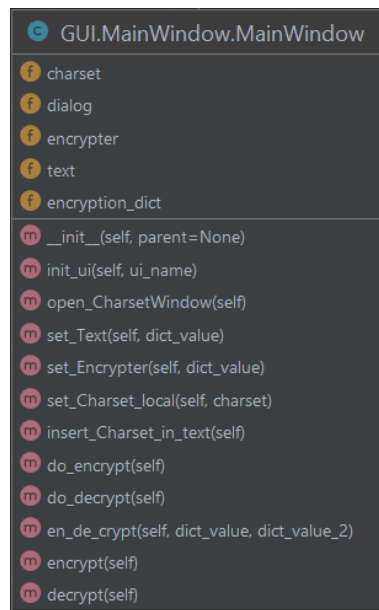
    # Menubar Settings
    ChangeCharset = self.actionChangeCharset
    ChangeCharset.triggered.connect(self.open_CharsetWindow)

    # Menubar Analysis
    self.actionKeyLength.triggered.connect(self.open_KasiskiTestWindow)

    self.encryption_dict = {
        "Vigenère": 0,
        "Caesar": 1,
        "Encrypt": 2,
        "Decrypt": 3
    }
```

5.1.3 Code Smells 3 Large Class

Der dritte Code Smells zeigt eine 'Large Class'. Dieser bezeichnet eine große Klasse die unter anderem zu viele Instanzvariablen, Methoden oder allgemein Codezeilen aufweist. Als Beispiel dafür ist in der unteren Abbildung die MainWindow-Klasse zu sehen.



Mit 130 Zeilen Code und 10 verschiedenen Funktionen ist der Umfang der Klasse etwas zu groß. Abhilfe könnte geschaffen werden indem man beispielsweise einige Funktionen die die Handhabung der Verschlüsselung übernehmen in eine neue View-Klasse auslagert o.Ä.

5.2 Angewendete Refactorings

Die beiden durchgeführten Refactorings beziehen sich auf andere Beispiele als oben besprochen.

5.2.1 Refactoring 1 Extract Method¹

Platzhalter

5.2.2 Refactoring 2 Replace Error Code with Exception²

Die Aufgabe dieses Refactorings ist es potentielle Fehler zu vermeiden die beim Fehlinterpretieren von Rückgabewerten o.Ä. auftreten können. Ziel ist es, dass ein Fehlerstatus explizit eine Exception auslöst. Dieses Refactoring wurde im Code für das Kasiski-Verfahren angewendet, dass je nach Länge des Textes keine Schlüssel-länge ermitteln kann.

```
def determine_key_length(self):
    try:
        self.KeyLengths_TE.setPlainText(str(guessKey(self.CipherTest_TE.toPlainText())))
    except ValueError:
        self.KeyLengths_TE.setPlainText("Unable to determine Length of Key from given Text")
    except:
        self.KeyLengths_TE.setPlainText("An Error occurred. Please try again with a different Text.")
```

¹Commit: 5eb9481d0b440b72332d2c3ed4208e68343f3ac4

²Commit: e7bfb622070be07df80dcd3635307d828a1e895

```
def find_key_length(self, ciphertext, seq_len, max_key_len):
    # find repeated sequences and their positions
    rsp = self._repeated_seq_pos(text=ciphertext, seq_len=seq_len)
    seq_spc = {}
    for seq, positions in rsp:
        seq_spc[seq] = self._get_spacings(positions)
    # calculate spacings between positions of each repeated
    # sequence and factor out spacings
    factor_lists = []
    for spacings in seq_spc.values():
        for space in spacings:
            factor_lists.append(self._get_factors(number=space))
    # get common factors by descending frequency,
    # which constitute candidate key lengths
    ckl = self._candidate_key_lengths(factor_lists=factor_lists, max_key_len=max_key_len)
    if len(ckl) == 0:
        raise ValueError
    return ckl
```

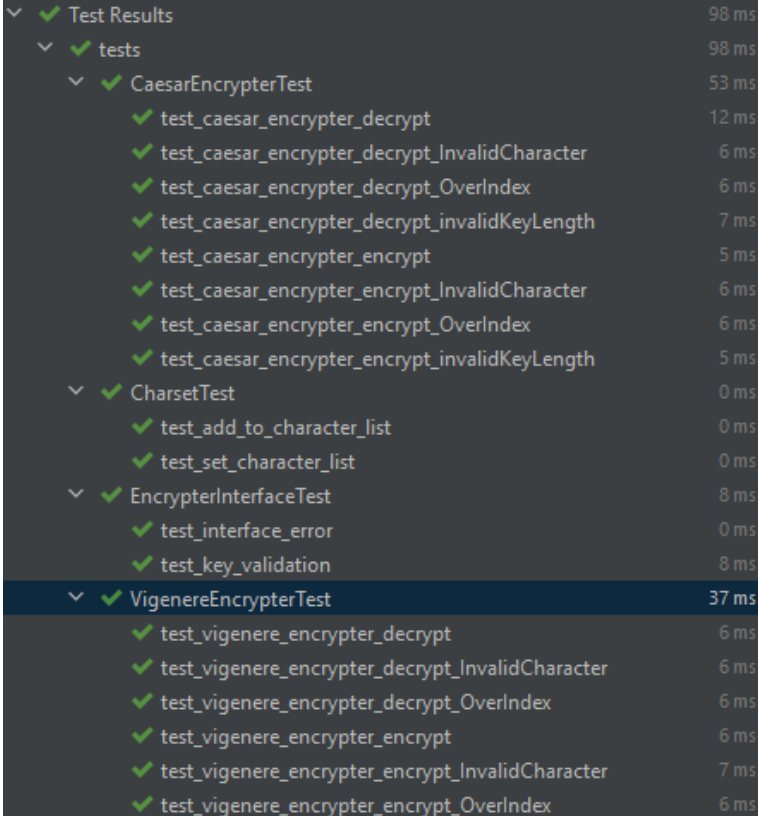
Für den Fall das die Funktion ' _candidate_key_lengths' eine leere Liste zurückgibt, wird statt einer Fehlermeldung ein ValueError ausgelöst. Dieser wird in der aufrufenden Funktion der Benutzeroberfläche abgefangen und entsprechend verarbeitet.

6 Unit Tests

Die Aufgabe von Unit Tests ist es einen jeweils möglichst kleinen Teil eines Systems zu testen. Dabei sollen nur relevante Teile des Systems im Test verwendet werden. Bestehen Abhängigkeiten zu anderen Teilen des Systems, so werden diese mit Stellvertreterobjekten besetzt. Ziel dieser Tests ist das Sicherstellen der Funktionalität der einzelnen Komponenten.

6.1 Verwendete Unit Tests und getesteter Code

Um den Umfang der Tests möglichst überschaubar zu halten wurden lediglich die Funktionen der Logik-Schicht getestet. Insbesondere die Verschlüsselung, aber auch die darin relevanten Charset- und Interface Klassen erhielten UnitTests.



Test Results	98 ms
tests	98 ms
CaesarEncrypterTest	53 ms
test_caesar_encrypter_decrypt	12 ms
test_caesar_encrypter_decrypt_InvalidCharacter	6 ms
test_caesar_encrypter_decrypt_OverIndex	6 ms
test_caesar_encrypter_decrypt_invalidKeyLength	7 ms
test_caesar_encrypter_encrypt	5 ms
test_caesar_encrypter_encrypt_InvalidCharacter	6 ms
test_caesar_encrypter_encrypt_OverIndex	6 ms
test_caesar_encrypter_encrypt_invalidKeyLength	5 ms
CharsetTest	0 ms
test_add_to_character_list	0 ms
test_set_character_list	0 ms
EncrypterInterfaceTest	8 ms
test_interface_error	0 ms
test_key_validation	8 ms
VigenereEncrypterTest	37 ms
test_vigenere_encrypter_decrypt	6 ms
test_vigenere_encrypter_decrypt_InvalidCharacter	6 ms
test_vigenere_encrypter_decrypt_OverIndex	6 ms
test_vigenere_encrypter_encrypt	6 ms
test_vigenere_encrypter_encrypt_InvalidCharacter	7 ms
test_vigenere_encrypter_encrypt_OverIndex	6 ms

6.1.1 Mocks

Für einige UnitTests ist es nicht möglich die einzelne Komponente abgekoppelt vom Rest zu testen. Um trotz dieser Abhängigkeiten zu anderen Klassen, Methoden, etc. die Funktionalität mit einem UnitTest sicherzustellen, werden Mocks eingesetzt. Deren Aufgabe ist es die benötigten Abhängigkeiten zu imitieren und die benötigten Schnittstellen zur Verfügung zu stellen. In diesem Projekt wurden Mocks auf zwei verschiedene Arten eingesetzt. In der unteren Abbildung ist die herkömmliche Art zu sehen.

```
def test_key_validation(self):
    # -----Mock
    self.mock = create_autospec(CaesarText)
    temp = string.ascii_uppercase
    character_list = []
    for i in range(len(temp)):
        character_list.append(temp[i])
    self.mock.get_eligible_characters.return_value = character_list
    # -----Test
    encrypter = CaesarEncrypter(self.mock)
    self.assertEqual(encrypter.validateKey("ABCDE"), "ABCDE")
    self.assertEqual(encrypter.validateKey("ABCDE"), "ABCD")
    self.assertEqual(encrypter.validateKey("ABCD%"), "ABCD")
```

Spezifisch für Python wird die 'create_autospec' Funktion auf die benötigte Klasse angewendet um ein Mock-Objekt zu erzeugen, welches die Eigenschaften der Klasse annimmt. Daraufhin wird für den Mock festgelegt was die später aufgerufenen Funktionen für einen Rückgabewert haben sollen. Zum Schluss wird der Mock im Test verwendet an der Stelle an der normalerweise die Ursprungs-klasse verwendet werden sollte.

Die zweite Art wie Mocks in den UnitTests in diesem Projekt verwendet werden ist in der nächsten Abbildung zu sehen.

```
def setUp(self):
    # -----Mock
    self.mock = create_autospec(VigenereText)
    temp = string.ascii_uppercase
    character_list = []
    for i in range(len(temp)):
        character_list.append(temp[i])
    self.mock.get_eligible_characters.return_value = character_list
```

Das Erzeugen und 'trainieren' des Mocks ist gleich wie im oberen Beispiel. Der Unterschied ist hierbei wie der Mock verwendet wird. Statt einen benötigten Mock für jeden Test individuell neu zu erzeugen und zu trainieren um ihn am Ende wieder zu verwerfen, kann man Tests mit ähnlichen Startbedingungen auch bündeln. Benötigen mehrere Tests die selben Mockobjekte, so bietet es sich an die 'setUp'-Methode zu verwenden. In einem Testbündel wird dadurch für jeden Test, vor der Ausführung dessen, diese Funktion einmal ausgeführt. Dadurch erhält man für jeden Test die selben Mocks, man spart sich aber Codeduplikate, also weniger potentielle Fehler und eine erhöhte Übersichtlichkeit.

6.1.2 Code Coverage³

Code Coverage beschreibt, welchen Anteil der Codezeilen im gesamten Projekt von Tests durchlaufen werden. Die Aufgabe dieses Wertes ist es einen groben Überblick

³Commit: 759310aada67666e973e748275b2739f1c93ace5

darüber zu geben, wie viel des Gesamtprojekts resistent gegenüber zukünftigen fehlerhaften Änderungen ist. Wichtig ist dabei aber nicht nur der absolute Prozentsatz sondern auch die Abdeckung der verschiedenen Branches einer Klasse, da nicht jede Stelle im Code gleich wichtig ist bzw. gleich oft aufgerufen wird. Mit den für dieses Projekt geschriebenen UnitTests wurde ein Coverage Report erzeugt, welcher in der unteren Abbildung zu sehen ist (Gefiltert nach eigenen Klassen).

	<i>statements</i>	<i>missing</i>	<i>excluded</i>	<i>coverage</i>
CaesarEncrypter.py	49	0	0	100%
EncrypterInterface.py	15	0	2	100%
VigenereEncrypter.py	50	0	0	100%
CaesarText.py	5	2	0	60%
EligibleCharacters.py	29	12	0	59%
Text.py	36	23	0	36%
VigenereText.py	5	2	0	60%
tests.py	0	0	175	100%
	189	39	177	79%

Wie zu erkennen ist werden nur die für die Logik verantwortlichen Klassen in der Coverage berücksichtigt. Weiterhin werden einige Zeilen vollständig in der Coverage verworfen. Dazu zählt beispielsweise auch die Klasse die für die Tests verantwortlich ist oder das EncrypterInterface in dem einige Zeilen lediglich eine 'NotImplemented'-Exception auslösen.

6.2 Anwendung der ATRIP-Regeln

Platzhalter