**Wrocław 01.10.2024**

**Cyberbezpieczeństwo**

**Laboratorium nr 2**

**kryptoanaliza Klasycznych algorytmów szyfrowania**

**Plan zajęć**

[I. Wstęp](#_Toc142213931)

[II. Zapoznanie się ze środowiskiem pracy - podstawowe narzędzia kryptoanalizy.](#_Toc142213932)

[III. Zadania:](#_Toc142213933)

[1. Algorytmy historyczne – właściwości.](#_Toc142213934)

[2. Analiza własności dostępnych algorytmów.](#_Toc142213935)

[3. Analiza dostarczonych plików.](#_Toc142213936)

[IV. Raport z przeprowadzonych eksperymentów.](#_Toc142213937)

# Wstęp

Zakres laboratorium obejmuje niektóre z podstawowych algorytmów kryptograficznych, które były kiedyś wykorzystywane w celu zapewnienia bezpiecznego sposobu przekazywania wiadomości między podmiotami. W tych algorytmach kryptograficznych przypisujemy liczby (lub) elementy algebraiczne do danego komunikatu wejściowego, który ma być przesyłany między dwoma bytami. Jeśli przypisane liczby (lub) wartości algebraiczne mają postać znaków z alfabetu danego języka, wówczas jest on traktowany jako zwykły tekst, który jest również nazywany tekstem jawnym. Ten zrozumiały tekst jawny jest konwertowany na niezrozumiałą formę zwaną tekstem zaszyfrowanym. Aby przekonwertować zrozumiały tekst jawny na niezrozumiały tekst zaszyfrowany, po stronie nadawcy używana jest funkcja szyfrowania. Podobnie, funkcja deszyfrująca jest używana po stronie odbiorcy, aby znaleźć zrozumiały tekst jawny dla niezrozumiałego tekstu zaszyfrowanego. Proces przekształcania zrozumiałego tekstu jawnego w niezrozumiały tekst zaszyfrowany i z powrotem na zrozumiały tekst jawny nazywa się przekształceniem kryptograficznym. W tym laboratorium omówiono historyczne algorytmy szyfrowania, takie jak algorytmy podstawieniowe, przestawieniowe (transpozycyjne) i algorytmy hybrydowe.

Obraz zawierający tekst

Opis wygenerowany automatycznie

**Entropia** to ilość informacji, jaką zawiera dana wiadomość lub zmienna. Entropia informacji jest czasami opisywana jako "liczba bitów potrzebnych do przekazania informacji". Z matematycznego punktu widzenia jest to suma prawdopodobieństw danego znaku i liczby wystąpień w obrębie danych. Okazuje się, że poziom entropii różnych typów danych jest różny. I tak, tekst w języku naturalnym będzie miał inny poziom entropii niż kod maszynowy, a ten z kolei będzie miał inny poziom entropii niż dane skompresowane. Entropia zaszyfrowanego ciągu danych, w zależności od zastosowanego algorytmu szyfrowania, może nie zmieniać się w ogóle lub znacząco wzrosnąć.

**Histogram** pokazuje względną częstość występowania każdego ze znaków w dokumencie. Pozwala nam to na porównanie częstości występowania różnych liter i może pomóc w określeniu charakteru informacji, a w przypadku dochodów nawet ujawnić niektóre reprezentacje znaków.

**Autokorelacja** służy do porównania tekstu z jego przesuniętą kopią, a nakładające się znaki są liczone. Okazuje się, że jeśli przesuwamy się o długość klucza (lub jego wielokrotność), to nakładających się znaków jest znacznie więcej. Znany również jako indeks koincydencji, jest to wskaźnik prawdopodobieństwa, że kiedy dwa teksty są porównywane litera po literze, dwa porównywane znaki będą takie same. Wykresy autokorelacji pokazują liczbę nakładających się znaków w zależności od przesunięcia tekstu. Na podstawie wykresów autokorelacji można określić długość użytego klucza szyfrującego.

**Okresowość** wyszukuje powtarzające się stałe wzorce w dokumencie, które zaczynają się od dowolnego punktu i trwają do końca dokumentu. Wzorzec ma dwie istotne cechy: przesunięcie i okres. Przesunięcie to pozycja początkowa cyklu, licząc od pierwszego znaku dokumentu, a okres to długość zestawu powtarzających się znaków.

**N-gram** dzieli wiadomość na bloki o rozmiarze N i sprawdza częstotliwość występowania fraz. Pozwala to na ocenę kluczy.

**Analiza losowości** dzieli wiadomość na bloki po 64 znaki, a następnie wyświetla zakres różnych liter dla bloku otaczającego każdą literę. Te wyniki są następnie wyświetlane użytkownikowi w formie wykresu liniowego, aby umożliwić mu zobaczenie, które części wiadomości mają wysoki poziom wariancji, a które mają stosunkowo niską wariancję. Celem jest określenie natury danych, które zostały zaszyfrowane. Analiza częstotliwościowa jest podstawowym narzędziem do łamania większości szyfrów klasycznych. W językach naturalnych, pewne litery alfabetu pojawiają się częściej niż inne. Badając te częstotliwości, można uzyskać pewne informacje o kluczu, który został użyty. Metoda ta jest bardzo skuteczna w przypadku klasycznych szyfrów, takich jak Cezar, Vigenère, itd. Jest ona jednak znacznie mniej skuteczna w przypadku nowoczesnych metod. W rzeczywistości, w przypadku nowoczesnych metod, najbardziej prawdopodobnym rezultatem jest uzyskanie pewnych podstawowych informacji o kluczu, ale nie otrzymanie klucza.

# Zapoznanie się ze środowiskiem pracy - podstawowe narzędzia kryptoanalizy.

* Program CrypTool 1.4.42 jest darmowym narzędziem rozwijanym przez uniwersytety w Siegen i Darmstadt dostępnym na stronie [www.cryptool.org](http://www.cryptool.org/) Program ten będzie podstawową platformą wykorzystywaną do realizacji kilku kolejnych ćwiczeń laboratoryjnych.
* Celem twórców CrypTool było i pozostaje nadal dostarczenie dobrej jakości pomocy dydaktycznej do ilustracji i demonstracji działania wybranych mechanizmów z dziedziny kryptologii.
* Oprócz samego programu realizującego kilkanaście popularnych algorytmów szyfrowania, pakiet zawiera bardzo dobrej jakości opracowanie na temat [kryptologii](https://www.cryptool.org/images/ctp/documents/CT-Book-en.pdf).
* Proszę zainstalować i przyjrzeć się możliwościom programu jeszcze przed zajęciami.
* Na pierwszych zajęciach będą wykorzystywane głównie narzędzia dostępne z zakładki *Szyfrowanie/Historyczne*

Uwaga!

Dla ćwiczeń dotyczących algorytmów historycznych, najlepiej korzystać z danych wejściowych w postaci ciągu znaków ograniczonych do alfabetu łacińskiego, bez znaków specjalnych i interpunkcyjnych.

Np.

DLACWICZENDOTYCZACYCHALGORYTMOWHISTORYCZNYCHNAJLEPIEJKORZYSTACZDANYCHWEJSCIOWYCHWPOSTACICIAGUZNAKOWOGRANICZONYCHDOALFABETULACINSKIEGOBEZ ZNAKOWSPECJALNYCHIINTERPUNKCYJNYCH

# Zadania:

## Algorytmy historyczne – właściwości.

Celem tego punktu jest zapoznanie się z podstawowymi algorytmami podstawieniowymi. W szczególności należy zwrócić uwagę na ich cechy wspólne jak i na elementy różnicujące.

**Zadania**

### Sprawdź dostępne algorytmy historyczne w zakładce Szyfruj/Odszyfruj -> Symetryczne (Klasyczne).

### Wybierz kilka z nich i sprawdź ich właściwości (tzn. przestrzeń kluczy, skutki wielokrotnego szyfrowania, czy istnieją tzw. klucze słabe np. dla Rot-13, Playfair, metody łamania szyfru). W raporcie należy umieścić właściwości wybranych algorytmów.

**Pytania/wnioski.**

### Co możemy powiedzieć o szyfrowaniu wielokrotnym w kontekście algorytmów historycznych (rozważ ich różne klasy)? Jak takie działanie wpływa na możliwość rozszyfrowania tekstu. Odpowiedź na to pytanie zilustruj wynikiem eksperymentu przeprowadzonego w Cryptool.

### Który z przetestowanych algorytmów może być uznany za silniejszy i dlaczego?

## Analiza własności dostępnych algorytmów.

### Porównaj wartości entropii tekstów jawnych dla różnych języków (angielski, polski, niemiecki, francuski, włoski, hiszpański, ...)

### Porównaj wartości entropii tekstu jawnego i tekstu zaszyfrowanego dla różnych algorytmów. W zadaniu należy zaszyfrować jeden tekst jawny za pomocą algorytmów wymienionych poniżej.

• Cezar,

• adfgvx,

• homofonów,

• permutacjnego,

• Vigenere (dla różnych długości kluczy),

• Hilla (dla różnych rozmiarów matryc),

### Porównaj histogramy tekstów jawnych dla wybranych 3 różnych języków (angielski, polski, niemiecki, francuski, włoski, hiszpański, ...).

### Porównaj histogramy jednego tekstu jawnego i zaszyfrowanego w zależności od algorytmu z punktu 2.

### Porównaj n-gramów (bi / tri / n) tekstów jawnych dla wybranych 3 różnych języków (angielski, polski, niemiecki, francuski, włoski, hiszpański, ...)

### Porównaj n-gramów (bi / tri / n) jednego tekstu jawnego i tekstu zaszyfrowanego w zależności od algorytmu z punktu 2.

### Przeanalizuj wartości autokorelacji tekstu zaszyfrowanego otrzymanego przez szyfrowanie algorytmami XOR i Vigenere dla różnych długości haseł.

**Pytania**.

### Jak zmieniają się obserwowane parametry w zadaniach od 1 do 7?

### W jaki sposób można wykorzystać narzędzia analizy tekstu dostępne w CrypTool do określenia algorytmu szyfrowania dla danego zaszyfrowanego tekstu?

### W jaki sposób można wykorzystać narzędzia analizy tekstu dostępne w programie CrypTool do ustalenia hasła używanego do szyfrowania?

## Analiza dostarczonych plików.

### Spróbuj rozpoznać, który algorytm szyfrowania został użyty, opierając się wyłącznie na analizie pliku z szyfrogramem. Wszystkie kryptogramy zostały utworzone na podstawie tego samego tekstu jawnego. Pliki 1\_X.txt, ..,

### Spróbuj odszyfrować ręcznie używając narzędzi analizy ( Kryptoanaliza -> Narzędzia analizy) np. n-gram, historgram, autokorelacja itd. lub używając opcji automatycznego łamania szyfru ( Kryptoanaliza -> Algorytmy historyczne-> Tylko szyfr) programy szyfrujące umieszczone w plikach 2\_x.txt, ...

### Rozpoznaj, jaki algorytm szyfrowania został użyty, a następnie spróbuj odszyfrować dostarczone programy z szyfrogramem. Pliki: 3\_x.txt, ...

**Pytania**.

### Od czego zależy siła algorytmu?

### W jaki sposób można zwiększyć siłę szyfrowania znanych szyfrów?

# Raport z przeprowadzonych eksperymentów.

Raport powinien mieć postać dokumentu elektronicznego, w którym zostanie zapisany przebieg realizowanych w trakcie zajęć eksperymentów (np. fragment treści tekstu jawnego, parametry algorytmu szyfrowania, fragment kryptogramu) oraz wnioski, spostrzeżenia, odpowiedzi na pytania umieszczone w instrukcji laboratoryjnej, itp.