Trabalho 1
Cifra de Vigenère

Davi Jesus de Almeida Paturi, 20/0016784

[1]Dep. Ciência da Computação – Universidade de Brasília (UnB)
CIC0201 - Segurança Computacional

davi.paturi@aluno.unb.br

Resumo. *A cifra de Vigenère é uma cifra polialfabética que utiliza uma chave repetida para cifrar e decifrar textos. Ela fornece uma camada adicional de segurança em comparação com cifras simples, mas ainda não é considerada segura para uso em criptografia moderna. Nesse trabalho abordaremos seu funcionamento e como quebrá-la.*

1. Introdução

A cifra de Vigenère é um método de criptografia clássico que foi inventado por Blaise de Vigenère no século XVI. Ela é uma cifra polialfabética, o que significa que utiliza várias tabelas de substituição (alfabetos) para cifrar o texto original, tornando-a mais resistente a ataques do que cifras simples, como a cifra de César.

A chave na cifra de Vigenère é uma palavra ou frase, chamada de chave de Vigenère, que é repetida ao longo do texto a ser cifrado. Cada letra da chave é associada a uma letra do alfabeto, e essa associação é usada para cifrar o texto original. A cifra de Vigenère funciona somando (ou subtraindo) as letras do texto original com as letras correspondentes da chave. Se a chave termina antes do texto, ela é simplesmente repetida até que tenha o mesmo comprimento do texto original.

Esse trabalho está dividido em 4 seções. A seção 2 apresenta as ferramentas e os metódos utilizados para o desenvolvimento desse trabalho. Na seção 3 está a implementação de modelos de cifragem e decifragem usando o cifra de Vigenère. A seção 4 expõe como atacar e decifrar essa cifra sem acesso a chave original. Na seção 5 está exposto os resultados e as vantagens e desvantagens dessa cifra.

2. Metodologia

A implementação dos algoritmos de cifragem e decifragem, bem como o algoritmo de ataque, foram feitos em Python. O código fonte das implementações estão disponíveis no Github [**?**].

3. Cifragem e Decifragem

O processo de cifragem na cifra de Vigenère envolve a utilização de uma chave de Vigenère para substituir as letras do texto original, uma por uma, de acordo com a chave. Como o deslocamento da cifra é baseado na chave e no alfabeto, conseguimos representar essa cifra algebricamente. Sendo C_i a letra criptografada, W_i a letra original e K_i a letra da chave de Vigenère, podemos representar a cifragem como:

$$C_i = W_i + K_i \; mod \; 26 \tag{1}$$

De forma análoga, conseguimos chegar a uma equação para a decifragem:

$$W_i = C_i - K_i + 26 \; mod \; 26 \tag{2}$$

A implementação em Python desses modelos de cifragem e decifragem segue abaixo.

```
def encrypt(text, key):
  text = text.upper()
  key = repeat_key(key, len(text))
  out = ""
  i = 0
  for c in text:
    if c in alphabet:
      out += alphabet[(ord(c) + ord(key[i])) % 26]
      i += 1
    else:
      out += c
  return out

def decrypt(text, key):
  text = text.upper()
  key = repeat_key(key, len(text))
  out = ""
  i = 0
  for c in text:
    if c in alphabet:
      out += alphabet[(ord(c) - ord(key[i]) % 26 + 26) % 26]
      i += 1
    else:
      out += c
  return out
```

4. Ataque

O ataque por frequência de letras na cifra de Vigenère é um método de quebra que se baseia na análise das frequências das letras do texto cifrado [Wikipédia 2023]. Ao dividir o texto cifrado em grupos correspondentes ao tamanho estimado da chave de Vigenère, o atacante examina as frequências das letras em cada grupo. Essa análise permite identificar padrões, pois cada grupo é essencialmente cifrado como uma cifra de substituição simples.

O objetivo principal do ataque é estimar o comprimento da chave de Vigenère e, em seguida, identificar as letras da chave. Para estimar o tamanho da chave utilizamos repetições de trigramas para tentar encontrar um padrão. A implementação em Python segue abaixo.

```
def key_size(text):
  text = valid_chars(text)
  interval = []
```

```
  for i in range(len(text) - 2):
    trigram = text[i] + text[i+1] + text[i+2]
    for j in range(i+1, len(text)-2):
      cand = text[j] + text[j+1] + text[j+2]
      if cand == trigram:
        interval.append(j - i)

  freq = {}
  for intv in set(interval):
    for i in range(2, 21):
      if intv % i == 0:
        freq[i] = freq.get(i, 0) + 1

  key_s = (0, 0)
  freq = dict(sorted(freq.items(), key = lambda i: i[1], reverse = True
   ))

  print("----- Possiveis tamanhos de chaves -----")
  print("--------- Tamanho | Quantidade ---------")
  for key, value in freq.items():
    if value >= key_s[1]:
      key_s = (key, value)
    print(f"{str(key).rjust(17, ' ')} | {value}")

  print(f"--------- Tamanho provavel = {key_s[0]} ---------")
  inp = input("Gostaria de trocar o tamanho da chave? (S/N)\n> ")
  if inp.upper() == "S":
    return int(input("Digite o tamanho de chave desejado:\n> "))

  return key_s[0]
```

Após encontrar o tamanho da chave, tentamos descobrir as letras da palavra chave baseando-se na frequência das letras na linguagem do texto. A letra escolhida é a com a frequência mais próxima a da frequência comum. A implementação do algoritmo de quebra segue abaixo.

```
def find_letter(prob, lang):
  lang = lang.lower()
  letter = ''
  total_diff = 999999999 # very high initial value

  for i in range(26):
    diff = sum(abs(prob[(i+j) % 26] - letter_freq[lang][j]) for j in
   range(26))
    if diff < total_diff:
      letter = alphabet[i]
      total_diff = diff

  return letter

def break_encryption(key, text, lang):
  text = valid_chars(text)
  keyword = ""
  for i in range(key):
    total = 0
    freq = {}
```

```
    prob = []
    for j in range(i, len(text), key):
      freq[text[j]] = freq.get(text[j], 0) + 1
      total += 1

    for c in alphabet:
      prob.append(freq.get(c, 0) / total * 100)

    keyword += find_letter(prob, lang)
  return keyword
```

5. Conclusão

Em conclusão, a cifra de Vigenère é um método de criptografia histórico que oferece uma camada adicional de segurança em comparação com cifras simples, como a cifra de César. No entanto, a cifra de Vigenère não é considerada segura para uso em criptografia moderna, pois pode ser quebrada com relativa facilidade usando técnicas como o ataque por frequência de letras, especialmente quando a chave é curta ou quando padrões na chave são detectáveis. Portanto, em ambientes onde a segurança é uma preocupação, é recomendável o uso de métodos de criptografia mais robustos e atualizados, como algoritmos de criptografia simétrica ou assimétrica, que oferecem maior proteção contra ataques sofisticados. A cifra de Vigenère é mais adequada como um exemplo histórico do desenvolvimento da criptografia do que como uma técnica de segurança moderna.

Referências

[Wikipédia 2023] Wikipédia (2023). Frequência de letras – wikipédia, a enciclopédia livre. https://pt.wikipedia.org/wiki/Frequłncia_de_letras. [Online; acessado em 29 de setembro de 2023].